AD-753 846

AN APPLICATION OF CEPSTRUM PROCESSING TO ACTIVE SONAR

Jack A. Shooter, et al

Texas University

Prepared for:

Naval Ship Systems Command

18 June 1971

DISTRIBUTED BY:

NTIS

National Technical Information Service
U. S. DEPARTMENT OF COMMERCE
5285 Port Royal Road, Springfield Va. 22151

AD753846

THE UNIVERSITY OF TEXAS AT AUSTIN

ARL-TR-71-25
18 June 1971

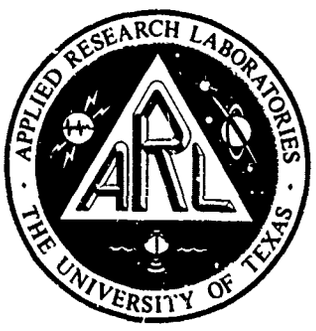
Copy No. 43

AN APPLICATION OF CEPSTRUM PROCESSING TO ACTIVE SONAR

Jack A. Shooter
Rodney E. Senterfitt

NAVAL SHIP SYSTEMS COMMAND
Contract N00024-70-C-1184
Proj. Ser. No. SF 11121103, Task 8614

Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
U S Department of Commerce
Springfield VA 22151



DDC
RECEIVED
JAN 12 1973
E

Approved for public release;
distribution unlimited.

ARL-TR-71-25
18 June 1971

AN APPLICATION OF CEPSTRUM PROCESSING TO ACTIVE SONAR

Jack A. Shooter
Rodney E. Senterfitt

NAVAL SHIP SYSTEMS COMMAND
Contract N00024 70-C 1184
Proj Sor. No. SF 11121103, Task 8614

APPROVED FOR PUBLIC
RELEASE, DISTRIBUTION
UNLIMITED

APPLIED RESEARCH LABORATORIES
THE UNIVERSITY OF TEXAS AT AUSTIN
AUSTIN, TEXAS 78712

I a

ABSTRACT

The use of cepstrum is investigated as a means of detecting a target and estimating target length in a low signal to reverberation ratio environment. A model experiment was performed in a fresh water lake using a scale model submarine for a target. The reverberation in the experiment was due mainly to the sound backscattered from the air-water interface. A major result is that the cepstrum output may be averaged from ping to ping to achieve a processing gain, resulting in a higher probability of detection. The estimation of target length is a classification clue and it is pointed out that if the receive data window covers the target then the cepstrum output is independent of epoch. That is, the cepstrum output is independent of where the target waveform is located within the received data window. Therefore, cepstrum is a capable tool for estimating target length. It is also shown that the cepstrum processor is complex and that averaging over several ping cycles is necessary in order to achieve the necessary processing gain over reverberation.

TABLE OF CONTENTS

ceding page blank

LIST OF FIGURES

Preceding page blank

LIST OF FIGURES (Cont'd)

I. INTRODUCTION

The Problem

One of the primary tools in antisubmarine warfare is the use of active sonar to detect and to classify underwater objects. Two of the problem areas in the use of sonar are the reverberation generated by the sound pulse and the complexity of most underwater targets. The problem of reverberation is concerned with detection. Frequently both targets and boundary surfaces are modeled as collections of discrete points. The significance of this type of model is that each discrete point backscatters a replica of the sound pulse and one cannot distinguish between points except on the basis of energy. That is, the statistical nature of each replica is the same but some points are better reflectors than others. Therefore a sonar is said to be reverberation limited. Once a target is detected then the problem becomes one of classification. It has been shown that the backscattered sound from simple targets does contain sufficient information to make fine distinctions between targets. For example, under laboratory conditions a likelihood ratio processor has been used to distinguish between hollow and solid spheres of slightly different sizes.[1] The success of the likelihood processor was dependent on the precision of knowing the arrival time (epoch time) of the backscattered sound relative to the transmit time. Under field conditions the epoch time is unknown, and is difficult to measure depending on the relative reverberation level. An additional problem with using a likelihood ratio processor to classify is the variability of a real complex target. That is, as the target changes

aspect relative to the sonar (the discrete point reflectors rearrange themselves), then the likelihood processor must change its model of the target in order to successfully classify the received sound.

An alternate procedure to using a likelihood processor or a matched filter[2] for classification is to utilize the physical parameters describing the target such as Doppler, target length, and aspect angle. These parameters are regarded as classification clues and it is the purpose of this report to show how cepstrum can be used to detect a target and to provide an estimate of target length in a low signal to reverberation ratio environment.

Cepstrum Background

Cepstrum, defined as the Fourier transformation of the log power spectrum, was first proposed[3] as a processor to detect the separation between overlapping waveforms as applied to seismic problems. However, its most successful application has been in the extraction of pitch in speech analysis.[4] When cepstrum is examined in detail, it is found to be a special case of homomorphic[5] filtering and it has been generalized by Oppenheim, et al.[6] Basically, a homomorphic process involves a multiplication or a convolution. A logarithm or logarithmic derivative is then used to separate the multiplicative process into an additive process. For example, suppose that it is desired to remove the effects of a low frequency function (trend) which is multiplied against a high frequency function. This would not be possible with an ordinary linear highpass filter, but one could take a logarithm of the total function, thus separating the two functions, highpass filter the log function, and then take the antilog of the filtered function[7] to restore the high frequency function. In the case of the backscattered sound from a submarine, there is a multiplicative process in the frequency domain. That is, suppose that a sound pulse $x(t)$ is projected into the water.

If the submarine consists of N point reflectors then a model of the backscattered sound from the submarine would be

$$r(t) = \sum_{n}^{N} a_n x(t-\tau_n) \quad , \tag{1}$$

where τ_n is the nth waveform arrival time (time after transmission) or epoch time, and a_n is the amplitude of the nth waveform. The Fourier transformation of Eq. (1) is

$$R(f) = \sum_{n}^{N} a_n X(f) e^{-i\omega\tau_n} \quad , \tag{2}$$

and the energy spectrum is

$$R^*(f)R(f) = X^*(f)X(f)\sum_{m}^{N}\sum_{n}^{N} a_n a_m e^{-i\omega(\tau_n-\tau_m)} \quad , \tag{3}$$

where R and X are the Fourier transforms of r and x respectively. There are two important features to note in regard to Eq. (3). The energy spectrum of the transmitted sound, $X^*(f)X(f)$, is multiplied against the double summation and can be removed (filtered) by a homomorphic process. The second feature is that Eq. (3) is a function only of frequency, f, and the time differences between epochs, $\tau_n - \tau_m$. If the energy spectrum X^*X of x(t) is uniformly constant in the frequency domain, then it can be easily removed by taking the log or log derivative of Eq. (3) and subtracting the mean value. If X^*X is unknown, then a suggested alternate procedure is to experimentally estimate it by projecting sound at a random surface (such as the air-water interface), receiving the backscattered sound, and estimating X^*X over 75 or more ping cycles. The estimation $\langle X^*X \rangle$ can be

accomplished this way very accurately as predicted by Ol'shevskii[8] and as shown experimentally by Plemons.[9] Once the estimate $\langle X^* X \rangle$ is obtained, it can be divided into Eq. (3), which effectively flattens the energy spectrum. The basic idea is to somehow remove the effect of $X^* X$ from Eq. (3) such that only the dependence on $\tau_n - \tau_m$ is left. A Fourier transformation of a filtered log energy spectrum is a cepstrum, which is almost the same as an autocorrelation. For example, the Fourier transformation of just an energy spectrum is an autocorrelation. The resulting cepstrum is a function of lag time or quefrency[3] and it will have peaks at the times $|\tau_n - \tau_m|$ relative to the cepstrum origin. Therefore, cepstrum will give an estimate of target length.

Contents

This report is empirical in nature. It should be noted that the data processing was performed with a general purpose digital computer. The second chapter discusses second order sampling (quantizing) as applied to bandlimited signals and includes a block diagram of the cepstrum processing used for the report. The third chapter discusses the cepstrum signal processing gain by treating cepstrum as a black box processor, and a comparison is made with other processors. The fourth chapter compares the range resolution of cepstrum with a replica correlator and shows how cepstrum can be used to estimate target length. The fifth chapter shows how cepstrum may be used as a detector in a low signal to reverberation environment and how it can localize a target in range. The sixth chapter examines the possibility of using a digital computer to compute cepstrum in real time. Chapter VII is a summary of conclusions and recommendations outlining the good and bad aspects of cepstrum.

II. PROCESSING TECHNIQUE

Sampling

The initial form for any sonar return is an electrical signal, which in this report is stored on analog tape. In order to implement computer techniques of data analysis, it is necessary to convert this analog signal to digital form, i.e., a time function represented by a sequence of values of that function taken at equally spaced intervals in time. In general this time spacing, Δt, must be such that the sampling frequency, $f_s = 1/\Delta t$, is twice the highest frequency present in order to prevent aliasing errors. Since f_s is constant throughout a sampling procedure, the notation $e(t) = e_i$ will be used where $t = i\Delta t$, $i = 0,1,2, \ldots$.

The data discussed in this report are restricted to signals whose energy is confined within a frequency band of width, W, and centered at ω_o, the carrier frequency. This prompts the consideration of quadrature sampling.[10] A signal satisfying these requirements may be represented by the equation

$$e(t) = x(t) \cos \omega_o t + y(t) \sin \omega_o t \quad , \tag{4}$$

where $x(t)$ and $y(t)$ are low frequency functions whose bandwidths are W centered at zero. They are referred to as the x and y quadrature components of the signal respectively. The minimum sampling rate for a signal of this nature is W on each quadrature component. It should be noted, however, that if a signal is sampled at W on each quadrature component, any energy at a frequency outside th and will be aliased

into the band. Figure 1 illustrates the quadrature sampling of a signal of slowly changing frequency whose original sampling rate was $4\omega_o/2\pi$. Each mark on the signal denotes a sample point. The x and y quadrature components are labeled as such. In this case the sampling rate is $\omega_o/2\pi$ on each quadrature component. The sampling rate is reduced to $W = \omega_o/(2\pi n)$ by considering only each nth x and y quadrature component pair. It becomes quite obvious that a great saving in computational time is made due to the decreased number of points necessary to represent a signal.

Cepstrum Processing

The cepstrum is defined as the inverse Fourier transform of the log power spectrum. Figure 2 is a flow chart of a computer program to calculate the cepstrum. The X array contains N x quadrature component samples of the signal and the Y array contains the corresponding N y quadrature component samples of the signal. The energy spectrum[11] is found directly from the Fourier transform of the signal. The Fourier transform is estimated digitally by use of the Cooley-Tukey[12] Fast Fourier Transform algorithm. The number of points transformed must be an integer power of two; thus if N does not satisfy this requirement, enough zeroes are added to the end of the X and Y arrays so that the number of points is an integer power of two; N is also modified to equal this number. Since cepstrum is a correlation process, N additional zeroes are placed at the end of each array to prevent a recycling[13] of the signal due to its harmonic nature. The energy spectrum of an artificially produced signal is shown in Fig. 3; since the sampling rate is W on each quadrature component, the frequency samples are independent. That is, the correlation between frequency components is near zero.[14] It is realized that for real data some aliasing will occur at this sampling rate, but it was found in this case that its effect is negligible when the data are sufficiently filtered. As mentioned in

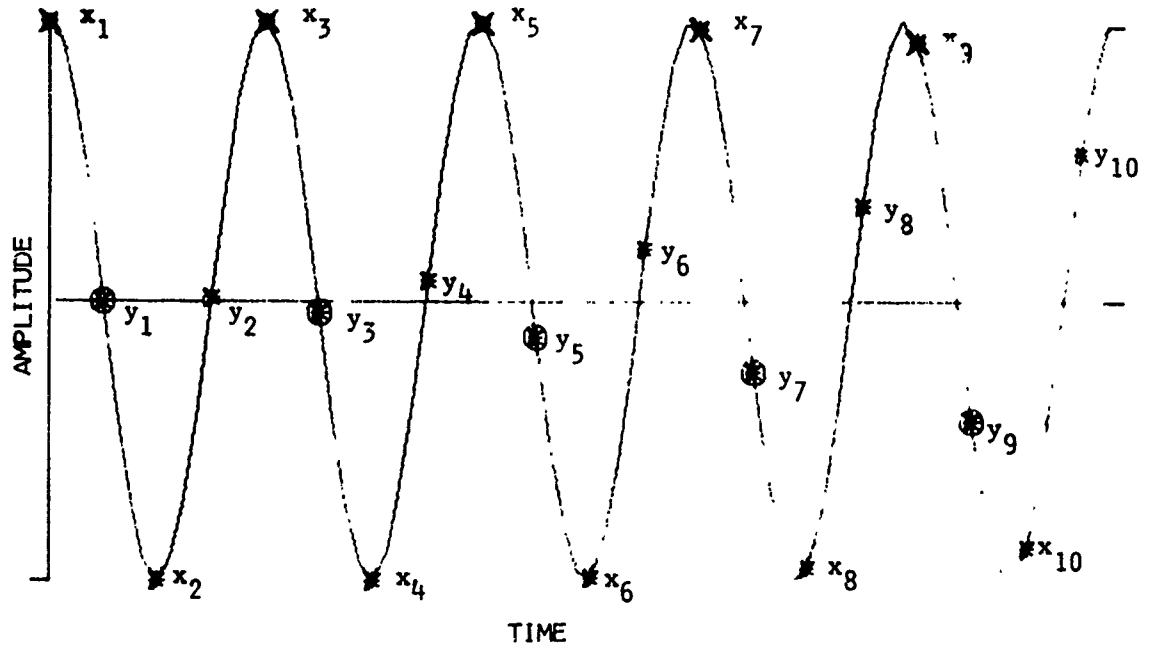


FIGURE 1

EXAMPLE OF QUADRATURE SAMPLING

AS-71-528

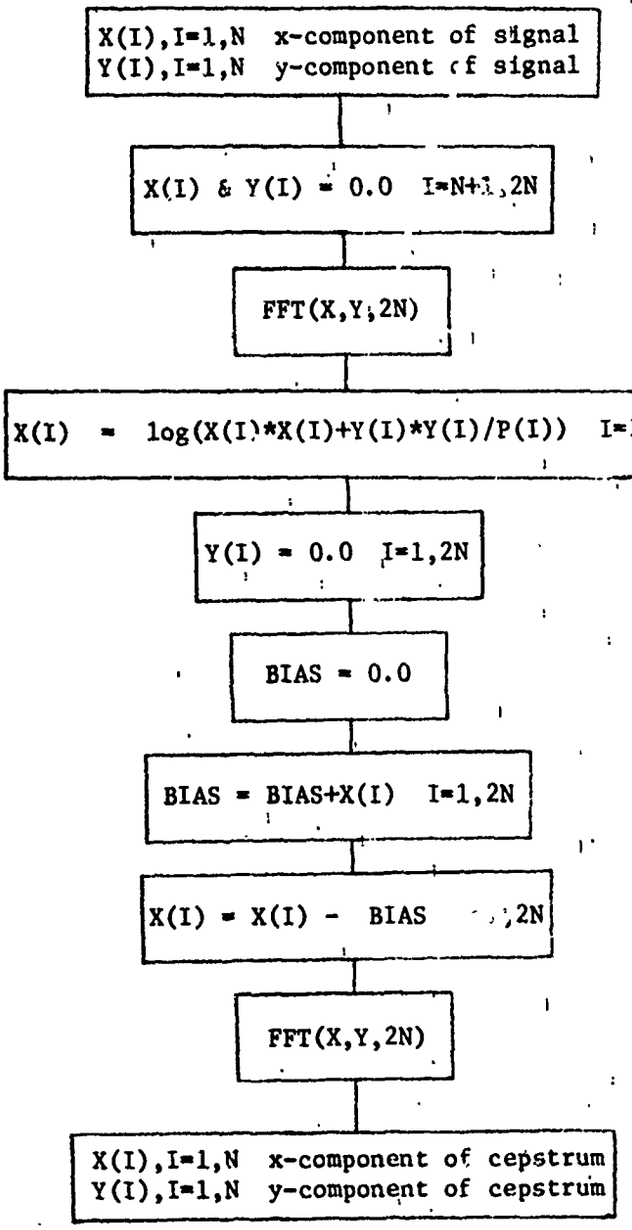
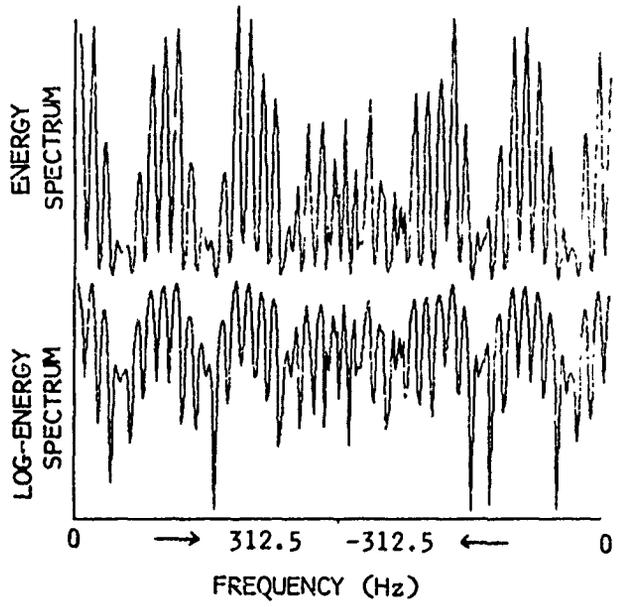


FIGURE 2

FLOW CHART OF CEPSTRUM COMPUTER PROGRAM

AS-71-529



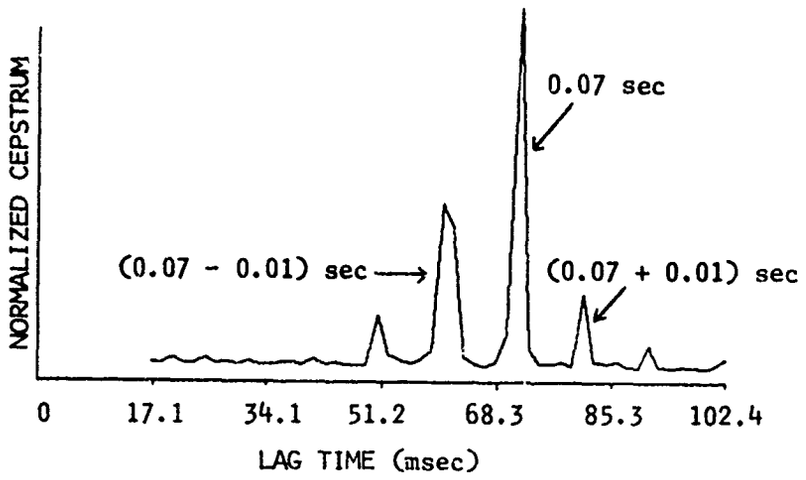


FIGURE 3

CEPSTRUM OF $Z(t) = \sin\omega t + \sin\omega(t - .01) + \sin\omega(t - .07)$

FM: 4687.5 → 5312.5 Hz TW: 64

$\tau_0 = 0$ sec, $\tau_1 = .01$ sec, $\tau_2 = .07$ sec

SAMPLING RATE = 624 Hz

NOTE: WITHOUT NOISE

the introduction, it is assumed that the effect of the frequency response of the entire electronic system may be found by calculating the estimated energy spectrum of a large number of signals. This average power P is divided into the energy spectrum of a single signal to remove the effects of the electronic system signal. The effect of the "flattened" signal energy is removed by extracting the mean log power form each point of the log power. The application of the Fast Fourier Transform to the X array, which contains the log power and the zeroed Y array, produces the x quadrature component of the cepstrum in the X array and its corresponding y quadrature component in the Y array. The cepstrum envelope is found by forming the modulus of the X and Y arrays. The log power and cepstrum envelope of an artificially produced signal along with its power spectrum is shown in Fig. 3. It will be noted that in a high signal-to-noise ratio situation the log process in the cepstrum will generate summation terms (e.g., (0.70 + 0.01) sec), which are not ordinarily visible in a low signal-to-noise situation.

III. PROCESSING GAIN

One method with which to evaluate the effectiveness of cepstrum as a processing technique is to compute its processing gain. Consider the cepstrum processor to be a black box with a certain signal-to-noise ratio on its input and a corresponding signal-to-noise ratio on its output. The processing gain, G, is defined here to be

$$G = \frac{\text{Peak}(s+n)_o^2 / N_o}{(S+N)_{in} / N_{in}} \quad , \tag{5}$$

where

$\text{Peak}(s+n)_o^2$ = peak signal plus noise out squared,

N_o = mean square noise out,

$(S+N)_{in}$ = mean square signal plus noise in,

N_{in} = mean square noise in.

This is a variation of the definition given by Stewart and Westerfield.[15] G is computed numerically as a function of S_{in}/N_{in} using artificially generated data for four different processors. Figure 4 shows the results for artificial data whose time-bandwidth product is 16. Notice that for a replica correlator, the crosscorrelation between a replica of the transmitted signal and the received signal, that the processing gain approaches twice the time-bandwidth, TW, product, which is the result predicted by Stewart and Westerfield for their definition of processing gain. The results for the polarity coincidence

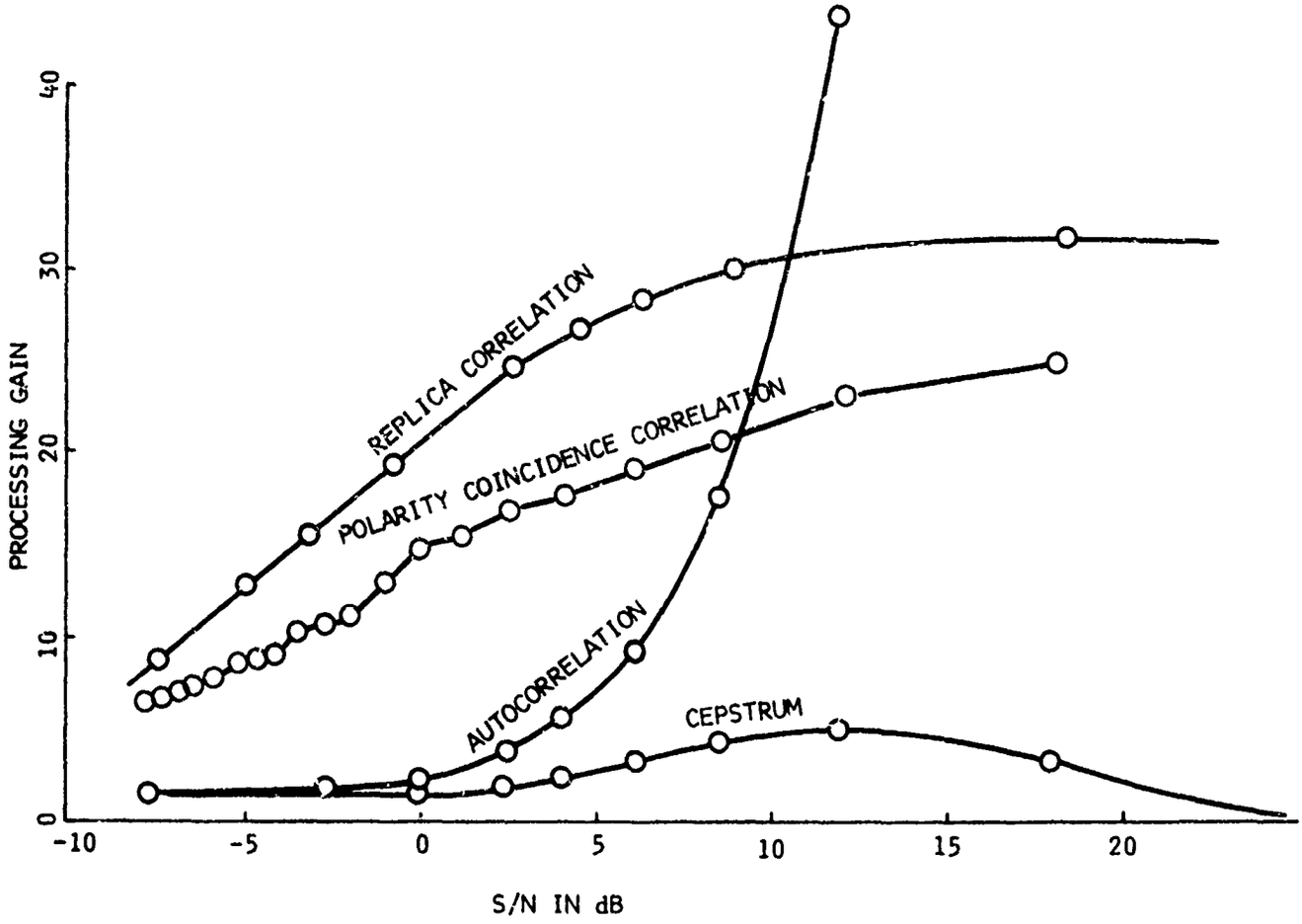


FIGURE 4

FOUR PROCESSING GAINS FOR
TW = 16

ARL - UT
AS-70-1442
RES - RFO
10 - 30 - 70

correlator are somewhat similar to those for the replica correlator. It is assumed that the degradation is due to nonavailability of amplitude information in the correlation.

The results for the autocorrelation and the cepstrum require an explanation of the procedure used in the calculation. The primary point of interest is to be able to detect the time difference between two received replicas of the transmitted signal superimposed in a noise background, as shown in Fig. 5a. The autocorrelation of Fig. 5a is shown in Fig. 5b, and the cepstrum of Fig. 5a is shown in Fig. 5c. The processing gain is computed by repeating these experiments 10 times and averaging to get the peak signal plus noise out squared. Typical results for the processes with noise only are shown in Fig. 6. The processing gain, G, for the autocorrelation grows as the square once the peak signal plus noise out is greater than the surrounding mean square noise out. The mean square output from the cepstrum is independent of the signal amplitude; consequently the processing gain goes to zero as S_{in}/N_{in} goes to infinity. To show this, consider a normally distributed, independent random variable, x, with a zero mean and a unity variance. Define a new variable, $y=\sigma x$. Let y be the noise input to the cepstrum processor.

$$N_{in} = \frac{1}{M}\sum_{i=1}^{M} y_1^2 = \frac{1}{M}\sum_{i=1}^{M} \sigma^2 x_i^2 = \sigma^2 \quad . \tag{6}$$

Without loss of generality it is assumed the variations in the signal-to-noise ratio are directly proportional to variations in the variance of the noise and not the target strength. If the Fourier transform of y is denoted F{y} the cepstrum, Cep, may be written

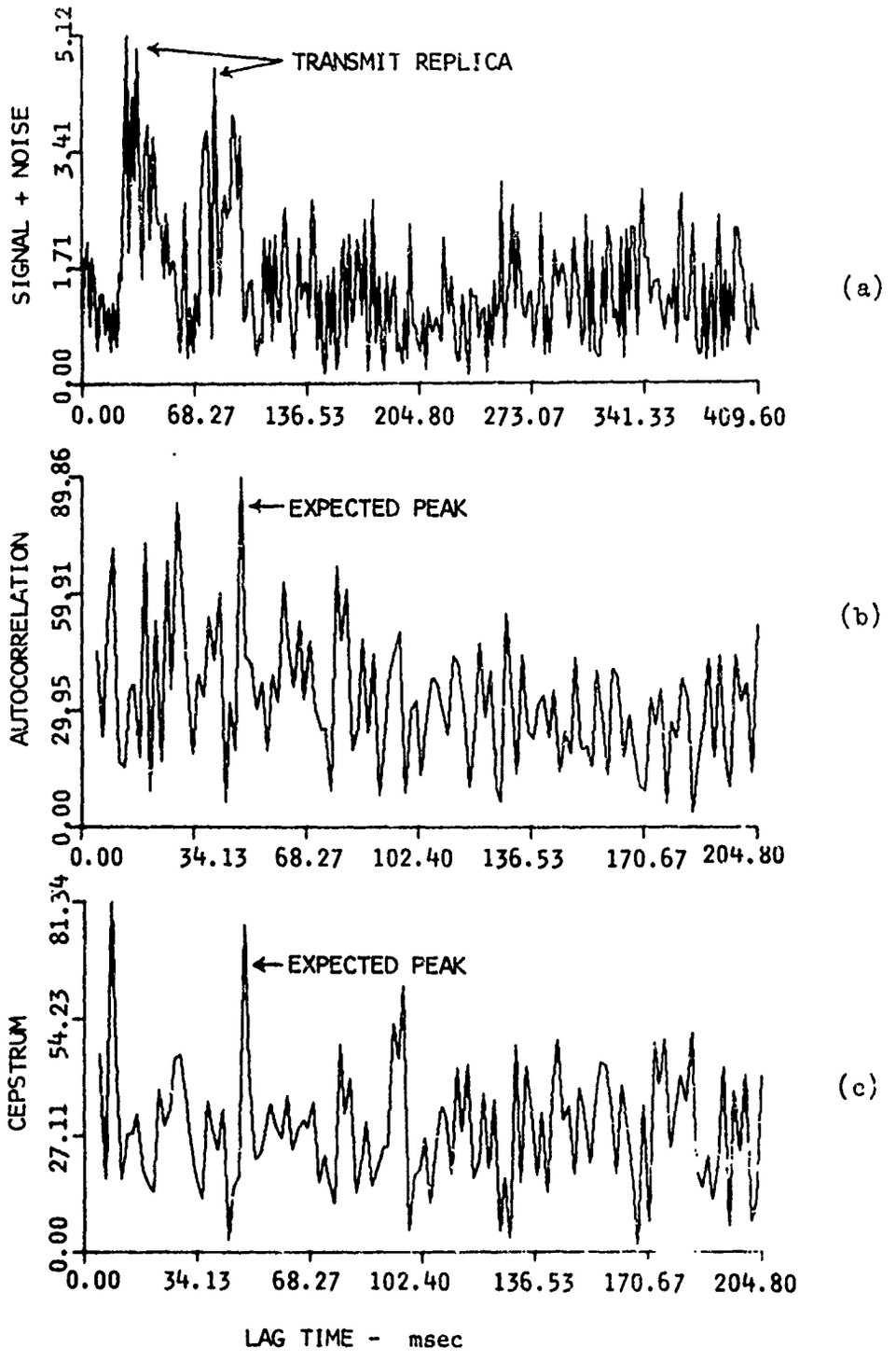


FIGURE 5

SIGNAL PROCESSING ON SIGNAL + NOISE

AS-70-1336

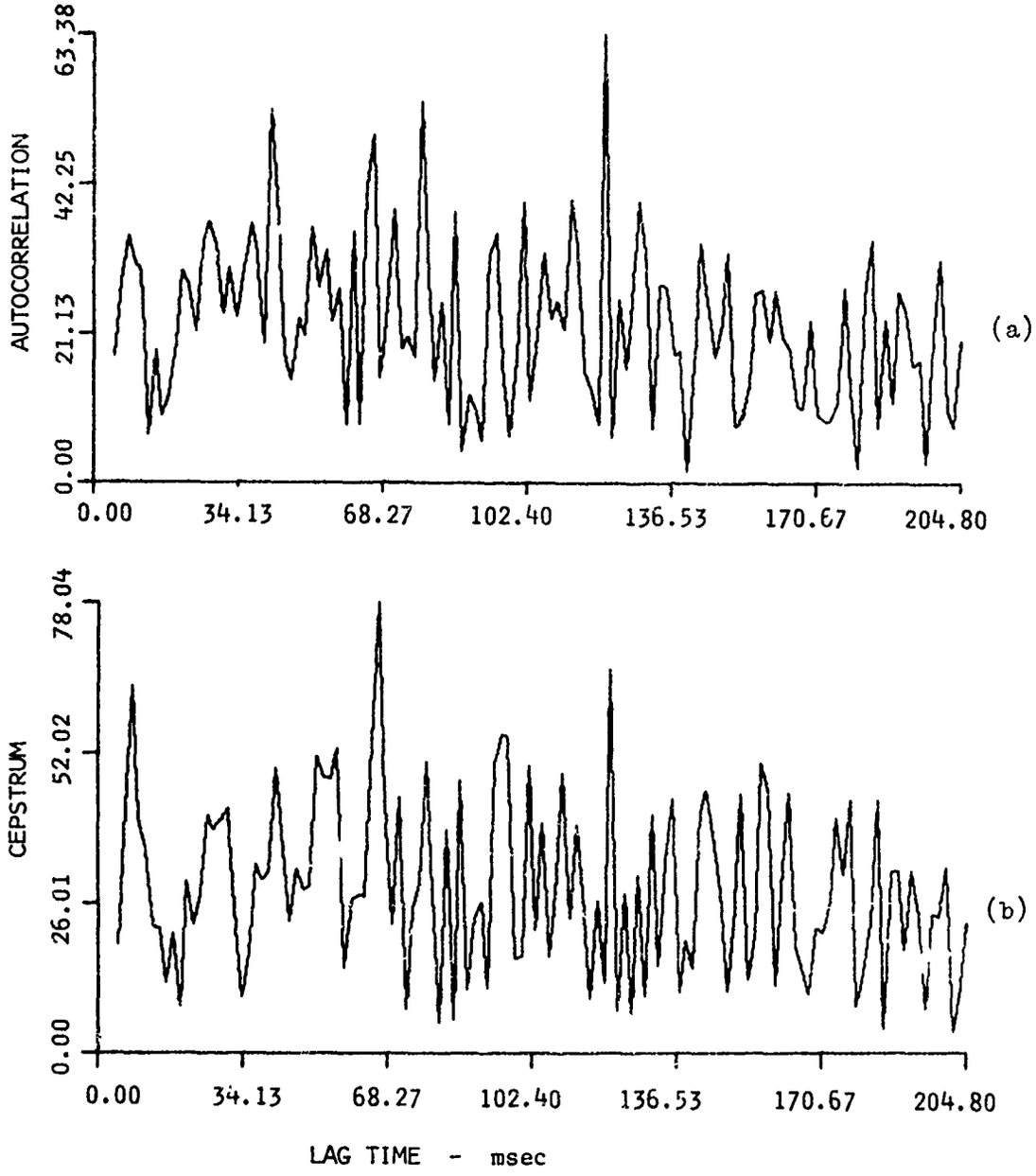


FIGURE 6

SIGNAL PROCESSING ON NOISE ONLY

AS-70-1337

$$\text{Cep} = \left|F\{\ell n(\,|F\{y\}|^2) - \text{mean}\}\right| \quad , \tag{7}$$

where

$$\text{mean} = \frac{1}{M}\sum_{i=1}^{M} \ell n(\,|F\{y\}|^2)_i \tag{8}$$

and

$$|F\{y\}|^2 = \sigma^2 |F\{x\}|^2 \quad . \tag{9}$$

Thus Eq. (7) reduces to

$$\text{Cep} = \left|F\left\{2\ell n|F\{x\}| - \frac{1}{M}\sum_{i=1}^{M} 2\ell n|F\{x\}|_i\right\}\right| \quad , \tag{10}$$

which is a constant with respect to σ^2, and therefore the mean square noise out is a constant, K. In the limiting case of large noise and small signal, the processing gain becomes

$$\lim_{S/N\to 0} \text{Est}\left[\frac{\text{Peak}(s+n)_o^2/N_o}{(S+N)_{in}/N_{in}}\right] = \frac{2K/K}{\sigma^2/\sigma^2} = 2 \quad . \tag{11}$$

where it is found that

$$\text{Est}(\text{Peak}(s+n)_o^2 = \text{Est}(N^2) = \int_0^\infty N^2 x_2^2(N)dN = 2K \tag{12}$$

for the dominating Gaussian noise, because N, which is defined in Eq. (12), is a chi-square variate with two degrees of freedom.[16]

For a large signal and comparatively low noise consider the special case of a received signal r, consisting of two superimposed replicas of the transmitted signal,

$$r(t) = ax(t) + ax(t-\tau) \quad . \tag{13}$$

Thus

$$\text{Cep} = \left|F\{\ell n 2a^2 + \ell n\left|F\{x\}\right|^2 + \ell n(1+\cos 2\omega\tau) - \text{mean}\}\right| \quad , \tag{14}$$

which shows once more that the cepstrum is independent of changes in the variance of the signal. The peak signal plus noise out is equal to a constant c; therefore

$$\lim_{S/N\to\infty}\left[\frac{\text{Peak}(s+n)^2_o/N_o}{(S+N)_{in}/N_{in}}\right] = \lim_{\sigma\to 0}\left[\frac{C/K}{\bar{a}^2/\sigma^2}\right] = 0 \tag{15}$$

where $\bar{a}^2$ denotes the mean square signal in. The results shown in Eqs. 11 and 15 agree very well with the experimental results shown in Fig. 4.

As would be expected the processing gain is a function of the transmit signal as is illustrated in Fig. 7, which presents the calculation of the processing gain of cepstrum for several different time-bandwidth products.

The conclusion is that the processing gain of cepstrum is low compared to other techniques; however, the processing gain may be increased by averaging the quadrature components before forming the envelope. It is not shown in this report, but a linear relationship was experimentally found between processing gain and averaging. That

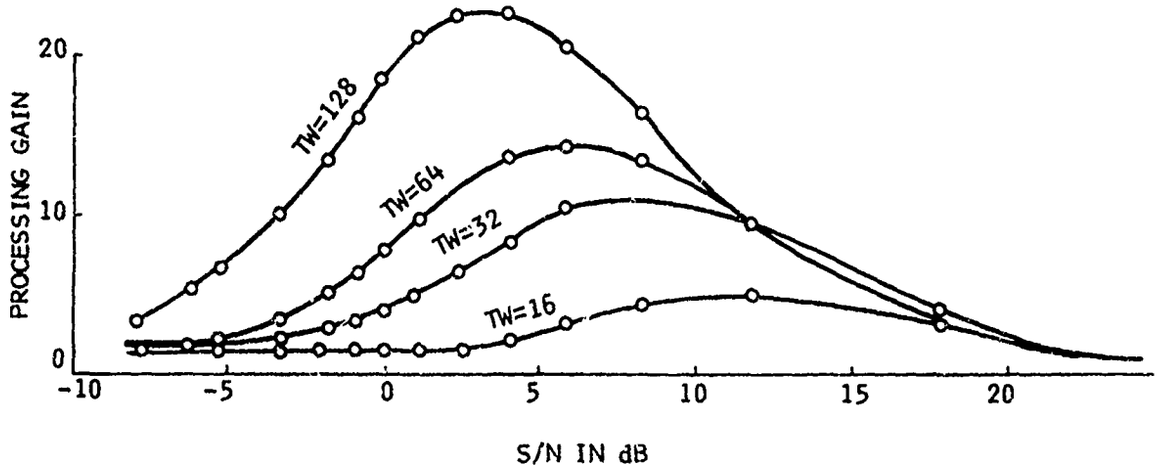


FIGURE 7

PROCESSING GAIN FOR CEPSTRUM

AS-70-1339

is, if N cepstrums are averaged then the processing gain is N times the curves shown in Fig. 7. From observing Fig. 7. it is found that it takes between 6 and 11 cepstrums to equal the processing gain of a replica correlator where the greater TW product requires the larger number of cepstrums for averaging.

IV. DETERMINATION OF TARGET LENGTH

Range Resolution

Assuming a "flat" energy spectrum, the range resolution of cepstrum is greater than that of replica correlation where the width of a peak from a replica correlator is $2/W$ measured from null to null. An exact relationship for the cepstrum resolution is not known due to its nonlinear nature. A comparison may be made, however, between replica correlation and cepstrum by using artificial data. The test signal contains two FM slides spaced Δt apart. Figure 8 gives an indication of the range resolution for the replica correlation for several different values of Δt. Figure 9 shows the cepstrum of the same signals. From Figs. 8 and 9 it is concluded that the range resolution of cepstrum is greater although there are greater side lobes.

Estimation of Target Length

It has been already noted that if a received waveform $r(t)$ consists of N superimposed waveforms $a_n x(t-\tau_n)$ from a target, that is,

$$r(t) = \sum_{n=1}^{N} a_n x(t-\tau_n) + \text{reverberation} \quad , \tag{1}$$

then the peaks in the cepstrum relating to the target will be located at the lag time $|\tau_n - \tau_m|$. If the target geometry relative to the sonar

Preceding page blank

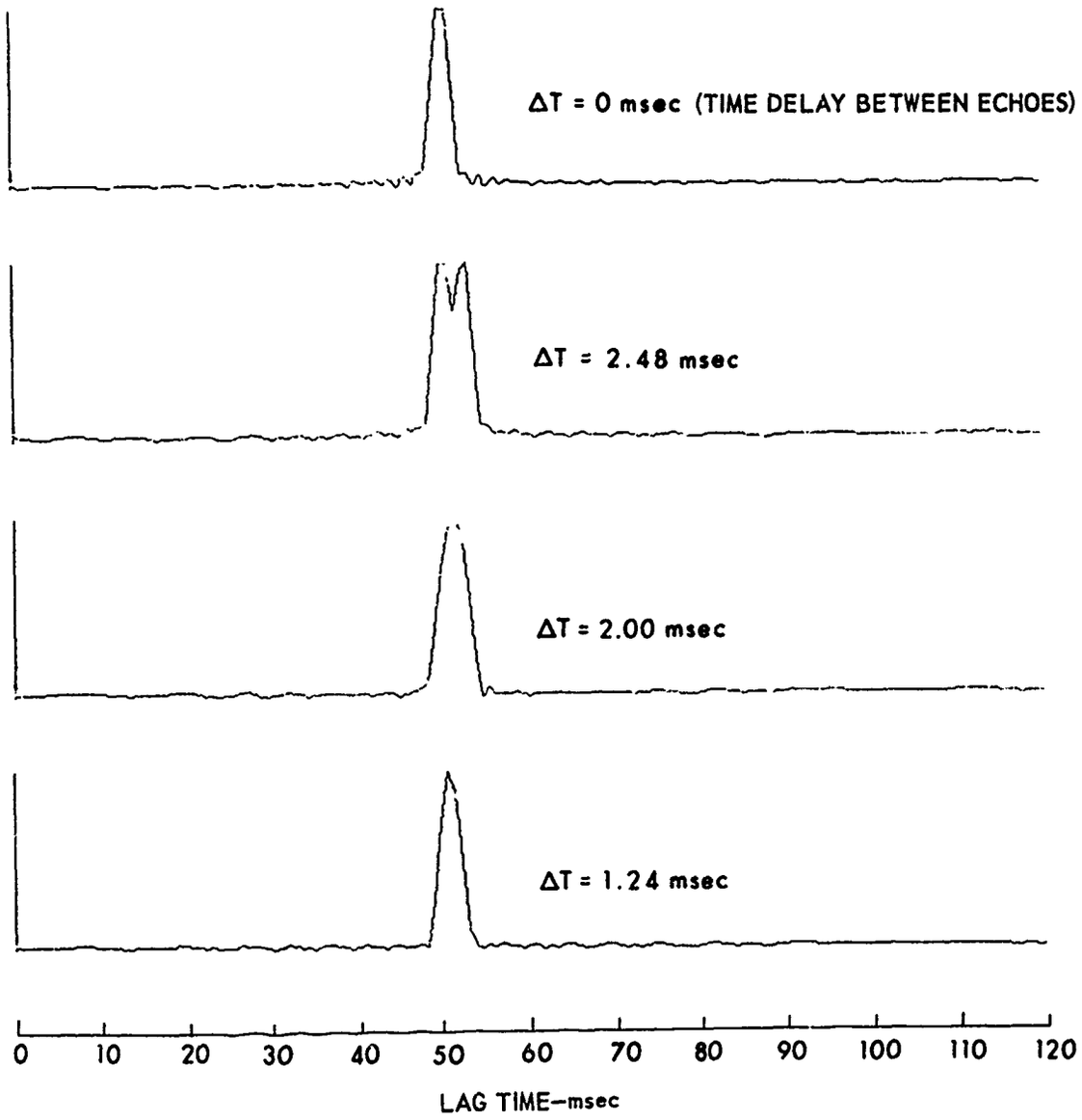


FIGURE 8
CROSSCORRELATION RESOLUTION TEST

LINEAR FM GAUSSIAN ENVELOPE 2:1
W = 625 Hz TW = 64 f_o = 5 kHz

ARL - UT
AS-69-230
JAS - JEW
2 - 17 - 69

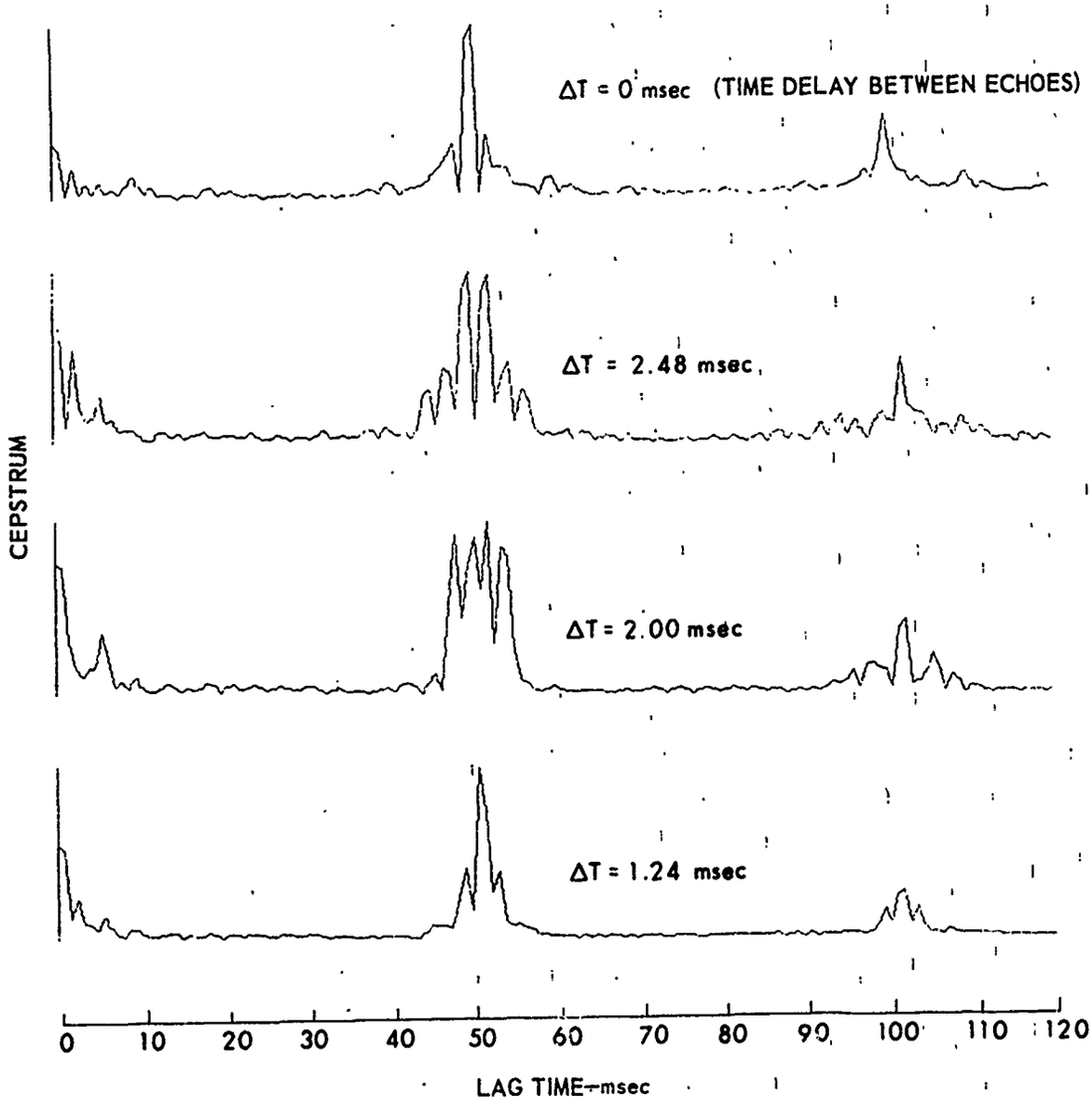


FIGURE 9
CEPSTRUM RESOLUTION TEST

LINEAR FM GAUSSIAN ENVELOPE 2:1
W = 625 Hz TW = 64 f_o = 5 kHz

ARL - UT
AS-69-231
JAS - JEW
2 - 17 - 69

is held constant then the τ_n, τ_m are constants and therefore the positions of the cepstrum peaks relating to the target will be constant while the cepstrum peaks relating to noise will vary in amplitude and position. As an illustration, a 2-point target was artificially generated, similar to those shown in Figs. 8 and 9. The 2-point target signal was superimposed in Gaussian noise at a 6 dB mean square signal-to-noise ratio. The resulting cepstrum was divided into 25 lag time gates and the experiment was repeated many times with differing sets of noise. The probability densities for each lag time gate are shown in Fig. 10, which is a joint probability density of lag time and amplitude. It is clear that there are consistently large peaks in the 6th lag time gate, which marks the target length or time separation between the two points. Note that the cepstrum noise appears to obey a Rayleigh law (which is expected since the original data is Gaussian). The amplitude probability density in the 6th lag time gate is a superposition of noise and signal. It is clear that, at least 50% of the time, the peaks in the 6th lag time gate are below the noise level. Therefore at the S/N level of 6 dB, cepstrum has a 50% probability of correctly estimating the target length. The accuracy of target length determination will be a function of S/N level, TW product (or processing gain), and how many cepstrums are averaged. The illustration of Fig. 10 does not show the effects of averaging.

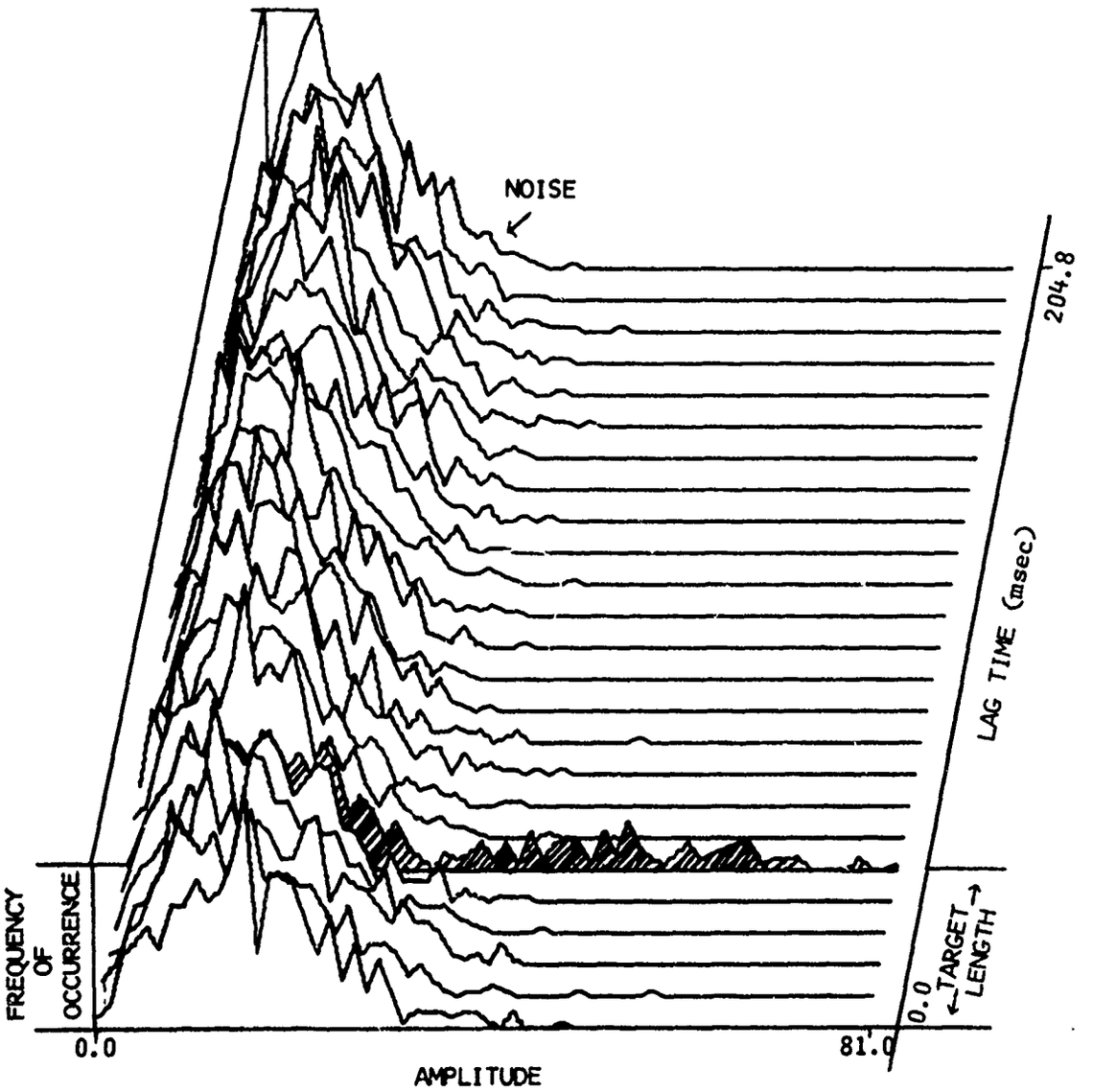


FIGURE 10

ESTIMATION OF A SIMPLE TARGET LENGTH USING A
JOINT PROBABILITY DENSITY (AMPLITUDE AND LAG TIME) OF CEPSTRUM
S/N = 6 dB

AS-71-531

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified

1 ORIGINATING ACTIVITY (Corporate author)	2a. REPORT SECURITY CLASSIFICATION
Applied Research Laboratories The University of Texas at Austin Austin, Texas 78712	UNCLASSIFIED 2b. GROUP ----

3 REPORT TITLE

AN APPLICATION OF CEPSTRUM PROCESSING TO ACTIVE SONAR

4 DESCRIPTIVE NOTES (Type of report and inclusive dates)

technical report

5 AUTHOR(S) (First name, middle initial, last name)

Jack A. Shooter and Rodney E. Senterfitt

6 REPORT DATE	7a. TOTAL NO OF PAGES	7b. NO OF REFS
18 June 1971	58	21

8a. CONTRACT OR GRANT NO	9a. ORIGINATOR'S REPORT NUMBER(S)
N00024-70-C-1184 b. PROJECT NO SF 11121103 c. Task 8614 d.	ARL-TR-71-25 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) —

10 DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited.

11 SUPPLEMENTARY NOTES	12 SPONSORING MILITARY ACTIVITY
	Naval Ship Systems Command Department of the Navy Washington, D. C. 20360

13 ABSTRACT

The use of cepstrum is investigated as a means of detecting a target and estimating target length in a low signal to reverberation ratio environment. A model experiment was performed in a fresh water lake using a scale model submarine for a target. The reverberation in the experiment was due mainly to the sound backscattered from the air-water interface. A major result is that the cepstrum output may be averaged from ping to ping to achieve a processing gain, resulting in a higher probability of detection. The estimation of target length is a classification clue and it is pointed out that if the receive data window covers the target then the cepstrum output is independent of epoch. That is, the cepstrum output is independent of where the target waveform is located within the received data window. Therefore, cepstrum is a capable tool for estimating target length. It is also shown that the cepstrum processor is complex and that averaging over several ping cycles is necessary in order to achieve the necessary processing gain over reverberation. (U)

DD FORM 1 NOV 68 1473 (PAGE 1)

S/N 0101-807-6801

UNCLASSIFIED

Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Cepstrum						
Covariance						
Quadrature sampling						
Ensemble statistics						
Reverberation						
Homomorphism						

V. CEPSTRUM AS A DETECTOR

Description of the Experiment

The effectiveness of cepstrum as a detector was tested by evaluating its ability to detect a target in a low signal-to-noise environment. A controlled experiment was conducted at the Applied Research Laboratories Lake Travis Test Station (LTTS). A rough sketch of the geometry of the experiment is shown in Fig. 11. Two hundred pings illustrated in Fig. 12 were recorded, digitized, and stored on magnetic tape. The transmit signal was a frequency modulated slide with the following characteristics:

Carrier frequency	114.87	kHz
Bandwidth	3.2	kHz
Pulselength	10	msec

The sampling was done in quadrature with a base frequency of 459.48 kHz.

Quadrature sampling rate (on each component)	6.4	kHz
Receive window length	160	msec
Receive window delay	70	msec
Number of data samples/component	1024	

Validity of the Ensemble

Before testing cepstrum as a detector it will be established that the 200 sonar returns constitute a valid ensemble.[17,18] That is,

Preceding page blank

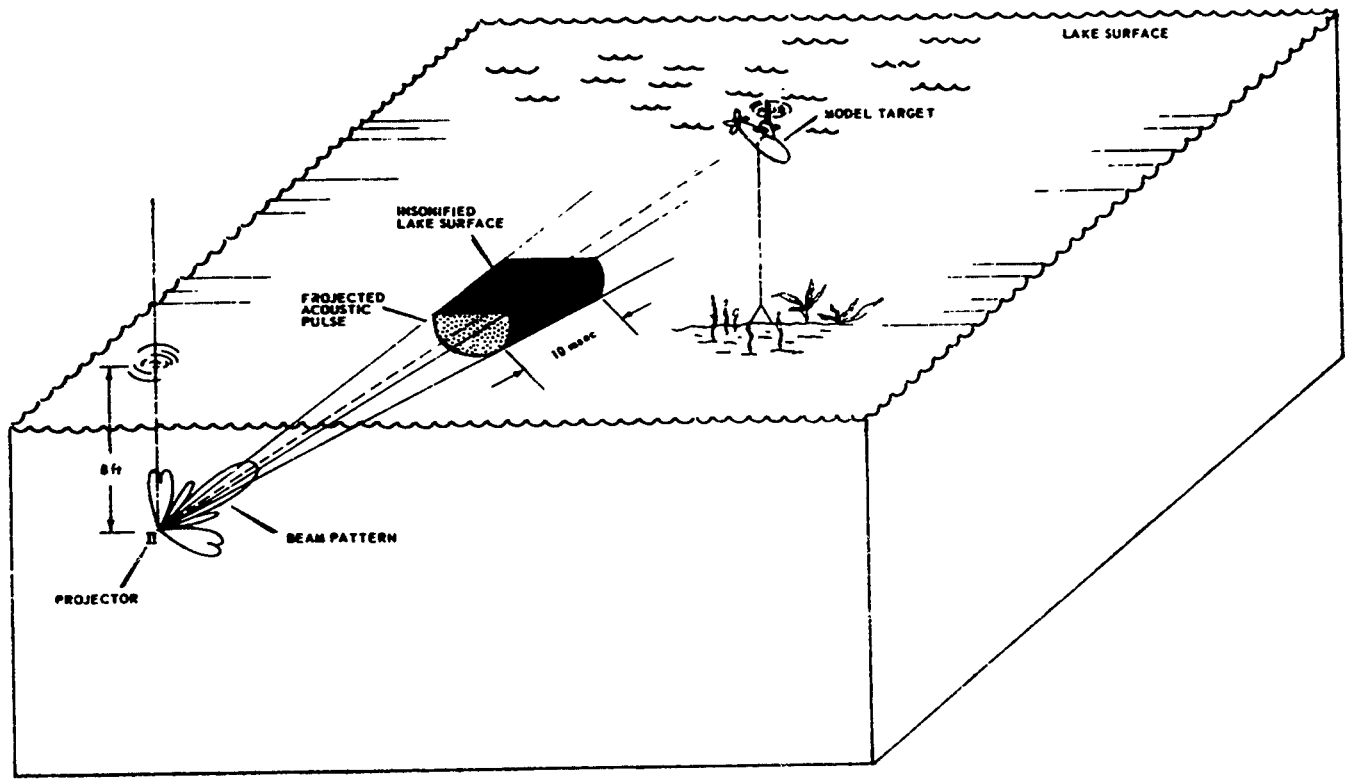


FIGURE 11

EXPERIMENTAL GEOMETRY AT THE
LAKE TRAVIS TEST STATION

ARL-UT
BS-71-520
JAS:ORS
8-12-71

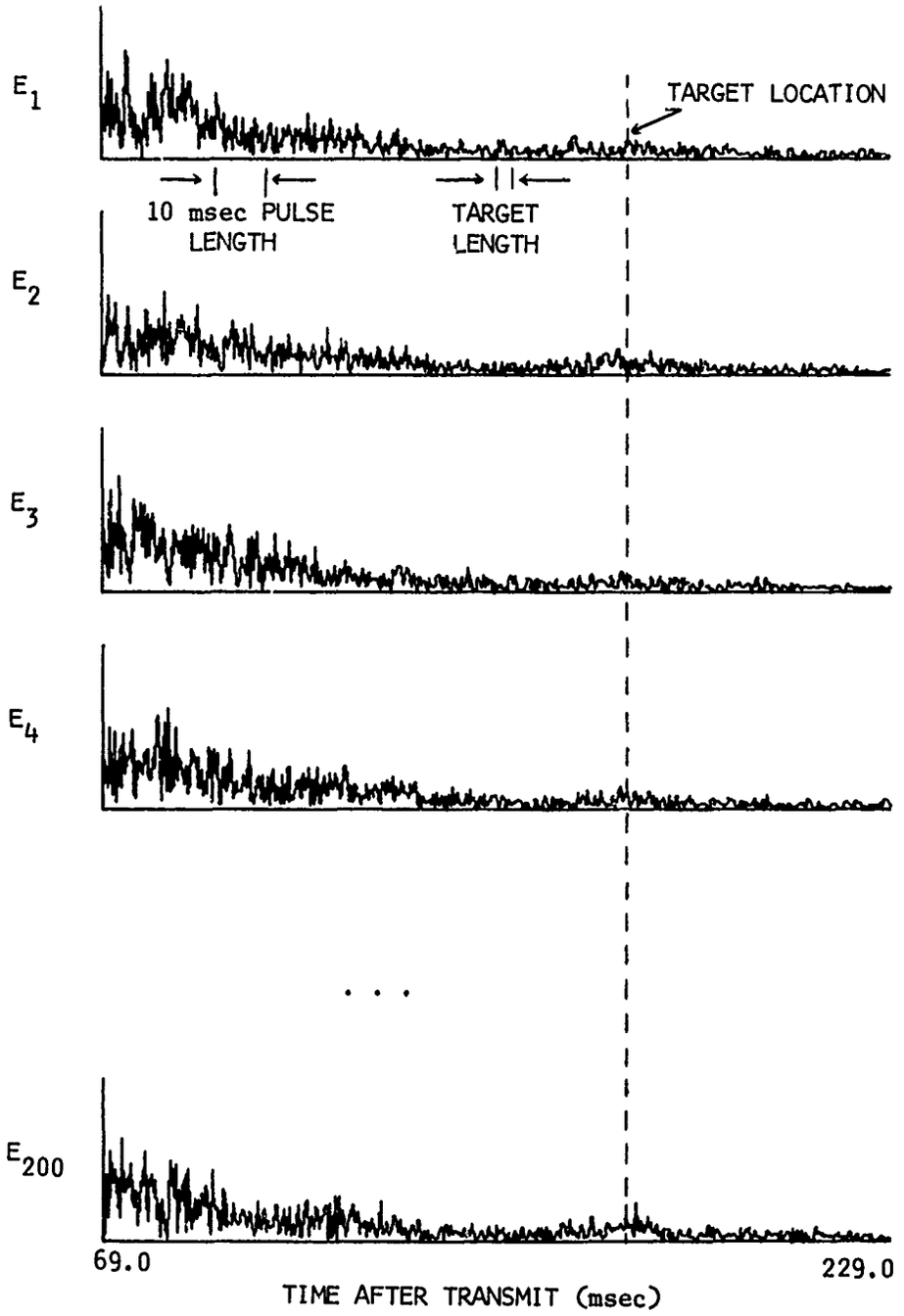


FIGURE 12

ENVELOPES OF SONAR RETURNS

AS-71-532

the data samples will be shown to be random, independent, and homogeneous. Actually, the ensemble will be shown to be valid as a function of range (time after transmission) because reverberation data are generally nonstationary in range. It is necessary to validate the ensemble because several cepstrums will be averaged together to obtain a processing gain and it would not be reasonable to average data that differ in their statistical parameters.

The test for homogeneity is done by two different hypothesis tests. The first, Kolmogorov-Smirnov two-sample test,[19] is done by comparing cumulative probability distributions taken at particular times, t_1, after transmission. For example, at time t_1 after transmission a sample (x quadrature component) is taken from each sonar return to form a random data set $(X_1(t_1), X_2(t_1), \ldots, X_{200}(t_1))$. The data set is subdivided into the first 100 numbers and the second 100 numbers. Then the test statistic, $Z(t_1)$, which is also a random variable, is measured by computing

$$Z(t_1) = \text{MAX} \; |F(X_i(t_1)) - F(X_j(t_1))| \qquad (16)$$

$$i = 1, 2, \ldots, 100$$

$$j = 101, 102, \ldots, 200 \quad ,$$

where the F's are the cumulative probability distributions of the experimental data. If the data are homogeneous, then the probability distribution of Z is known and a threshold may be set for a given confidence level. The result is shown in Fig. 13 where $Z(t_1)$ is plotted as a function of t, or range, and the threshold levels for $\alpha = 0.05$ and $\alpha = 0.50$ are shown as horizontal lines. The interpretation of Fig. 13 is that if the data are homogeneous at time t_1 then there is a 95% probability that $Z(t_1)$ will lie below the $\alpha = 0.05$ line or equivalently there is a 50% probability that $Z(t_1)$ will lie below the $\alpha = 0.50$ line. It will also be noted that Z is nonparametric, i.e.,

not dependent on the sonar return intensity, because the cumulative probabilities are always normalized to one. The second test for homogeneity, Wald-Wolfowitz two-sample runs test,[19] uses the same data, where the first 100 numbers are put into a set A and the second 100 numbers are put into a set B. The test statistic $Z(t_1)$ is found by combining the sets AB, sorting the data into a nondecreasing sequence, and then counting the number of runs N_A of A and the number of runs N_B of B. Then

$$Z = \frac{r - \mu_r}{\sigma_r} \quad , \tag{17}$$

where

$$r = N_A + N_B$$

$$\mu_r = \frac{2N_A N_B}{N_A + N_B} + 1 \quad ,$$

$$\sigma_r = \sqrt{\frac{2N_A N_B(2N_A N_B - N_A - N_B)}{(N_A + N_B)^2 (N_A + N_B - 1)}} \quad ,$$

and the result is shown in Fig. 14 along with the threshold lines. It is interesting to compare Figs. 13 and 14 and to note that significant events, i.e., points above the $\alpha = 0.05$ line, do not coincide, but the number of significant events in each figure are relatively the same. It will be recalled that even if the data are homogeneous there is a 5% probability that the test statistic Z at any point in time will lie

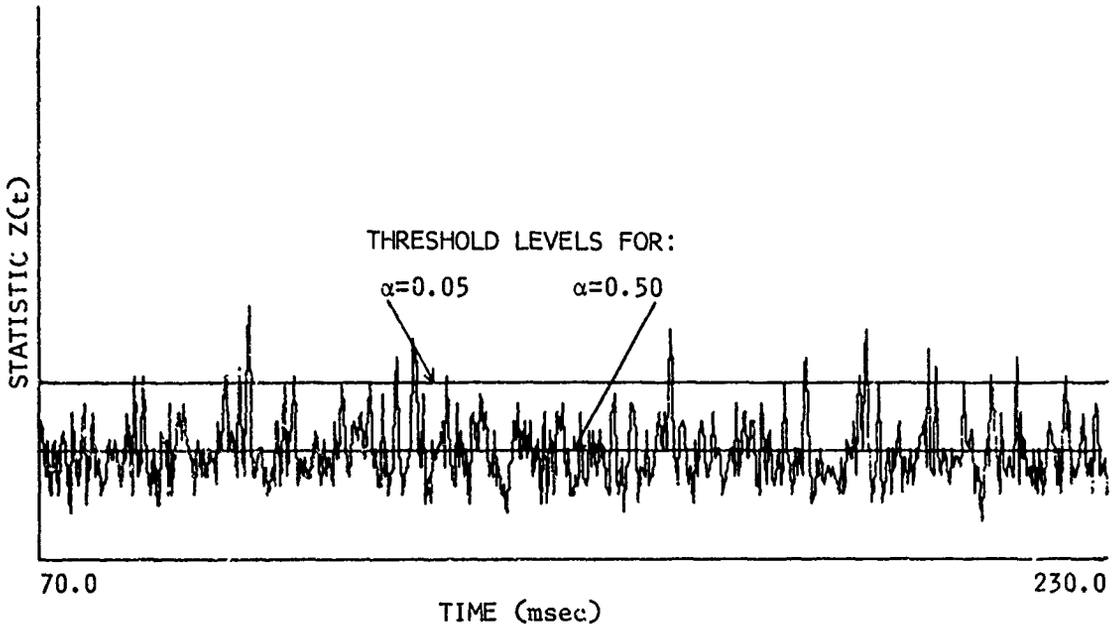


FIGURE 13

KOLMOGOROV-SMIRNOV TWO-SAMPLE TEST FOR HOMOGENEITY

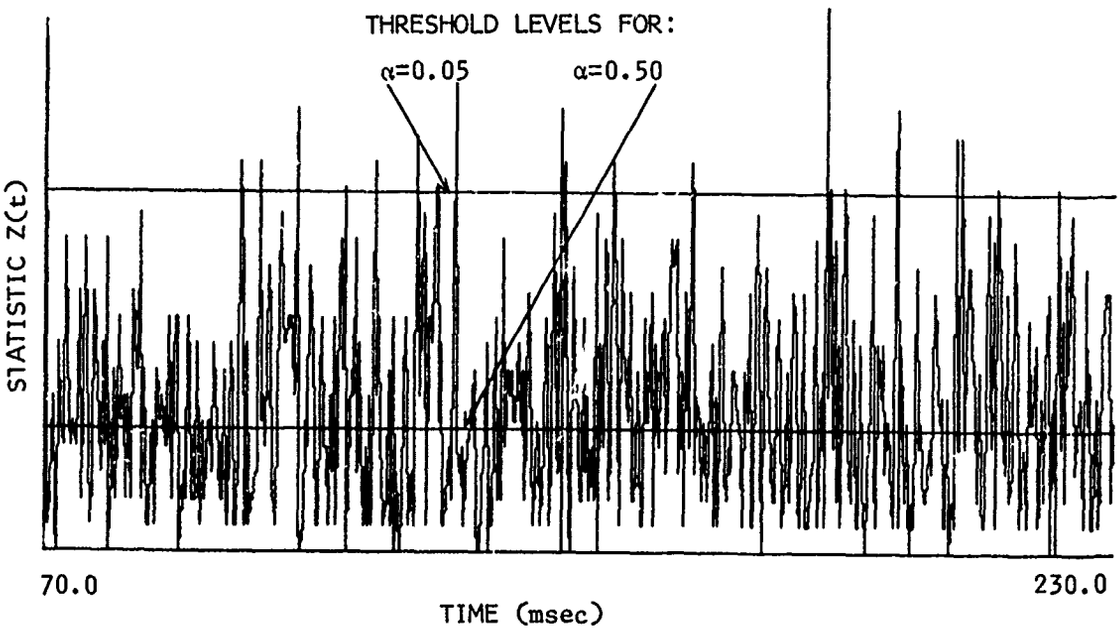


FIGURE 14

WALD-WOLFOWITZ TWO-SAMPLE RUNS TEST FOR HOMOGENEITY

AS-71-533

above the $\alpha = 0.05$ line. To complete the test for homogeneity it would be necessary to subdivide the data again and retest it; however, from the results it is concluded that these data are homogeneous.

It is also necessary to test the data for independence, which will be done two ways. The covariance matrix, which is discussed later, is one way, and another is to use a one-sample runs test.[19] The same data as used in Figs. 13 and 14 were treated as positive or negative numbers and runs of pluses and minuses were counted. The test statistic Z is computed in a similar fashion as in Eq. (17) and the result is plotted in Fig. 15. It is concluded from Fig. 15 and the covariance matrix that the data are independent

In addition, a test for normality, which is unnecessary to establish the validity of the ensemble, is done on the same data. A Kolmogorov-Smirnov one-sample test is used which is similar to the technique described by Eq. (16) except that one of the cumulative distributions is theoretical, with its mean and variance computed from the experimental data. The test statistic is plotted in Fig. 16 and it is noted that there are no significant events above the $\alpha = 0.05$ line. The conclusion is that the data are normally distributed as a function of range, and the test is very conservative because the experimental means and variances were used to compute the theoretical distributions.

To establish the independence of the data and to obtain the intensity of the reverberation and the backscattered sound from the target, consider the covariance matrix of the ensemble just described. Since the reverberation data are nonstationary, ensemble averaging is used. The covariance matrix, $C(t_1,t_2)$, is calculated by the formula[20]

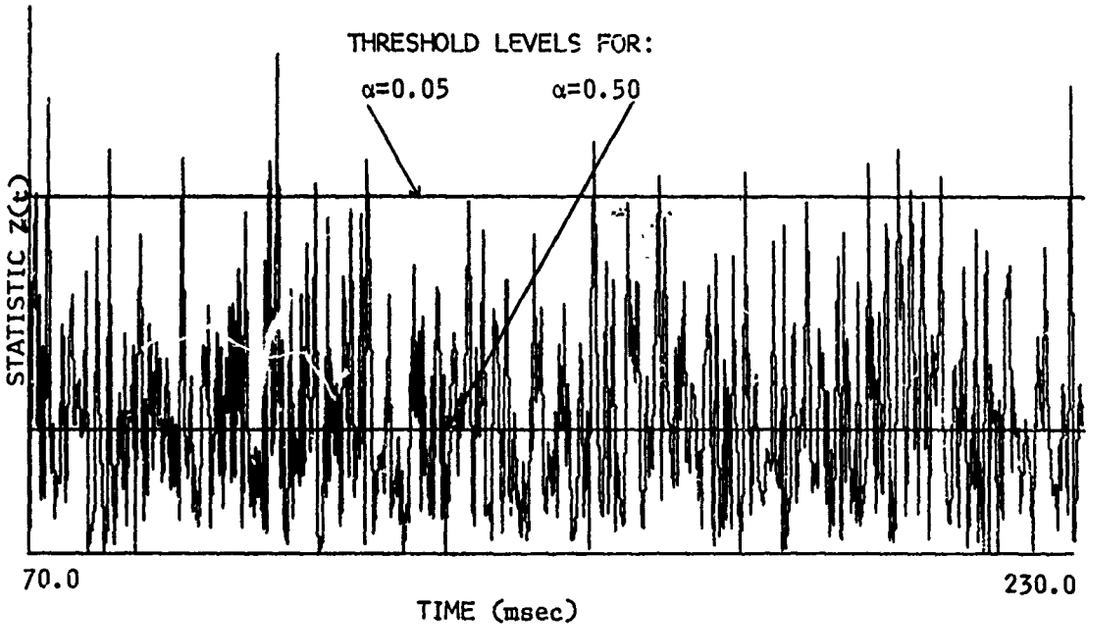


FIGURE 15

ONE-SAMPLE RUNS TEST FOR INDEPENDENCE

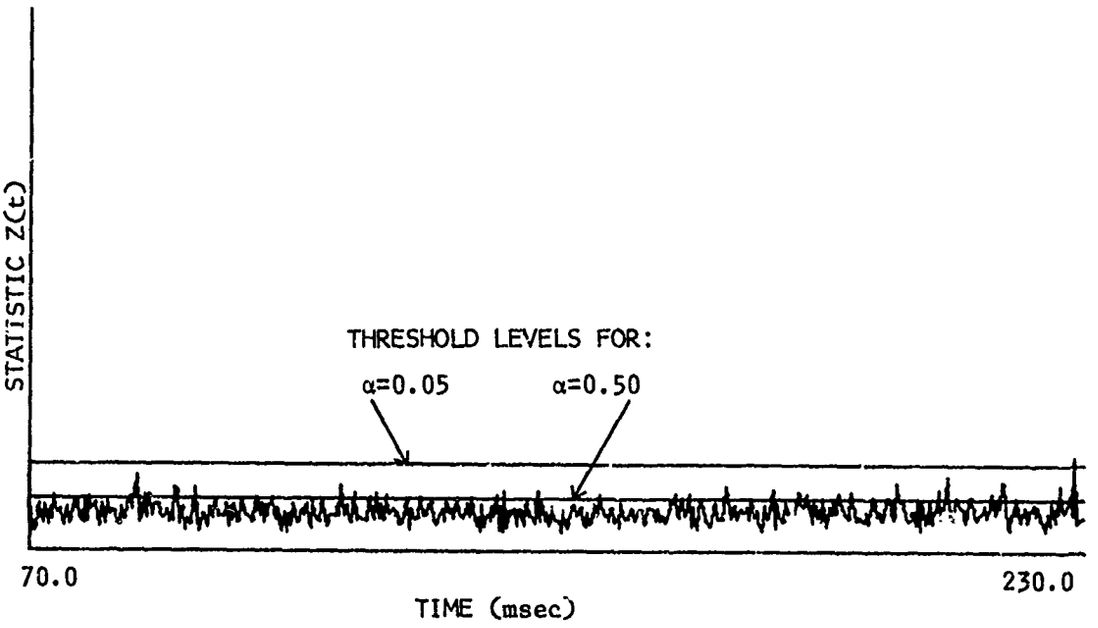


FIGURE 16

KOLMOGOROV-SMIRNOV ONE-SAMPLE TEST FOR NORMALITY

AS-71-534

$$C_x(t_1,t_2) = \frac{1}{2}\left[X(t_1)X(t_2) + Y(t_1)Y(t_2)\right]$$
$$C_y(t_1,t_2) = \frac{1}{2}\left[Y(t_1)X(t_2) - X(t_1)Y(t_2)\right] \qquad (18)$$

where $C_x(t_1,t_2)$ and $C_y(t_1,t_2)$ are the x and y quadrature components of the autocovariance, respectively, and the envelope, $E(t_1,t_2)$, which is plotted in Fig. 17, is found by the equation

$$E(t_1,t_2) = \sqrt{C_x(t_1,t_2)^2 + C_y(t_1,t_2)^2} \quad . \qquad (19)$$

Also shown in Fig. 17 is an enlarged section of the covariance matrix which covers the target region. As an aside it should be noted that the covariance matrix is a bandlimited two-dimensional high frequency surface. The plot in Fig. 17 is the envelope of that surface showing a strip along the main diagonal of the covariance matrix. The width (null to null) of the ridge running along the main diagonal is 2/W, where W = 3.2 kHz. Note that the ratio between the target intensity and the intensity of the immediately surrounding reverberation is in the range 1.5 to 2.0. Since cepstrum is independent of relative changes in intensity, the high intensity reverberation at the beginning of the signal will have a negligible effect on the processing. A plot similar to Fig. 17 is shown in Fig. 18 for the normalized covariance matrix. Notice that the width of the major lobe (diagonal) is fairly constant except in the region of the target, where a broadening effect is present due to the finite length of the target.

It has been determined that an increase in processing gain is obtained by averaging the x and y quadrature components of the cepstrums of several homogeneous pings. The effect is to average out the noise which is random in amplitude and phase, and to have a much

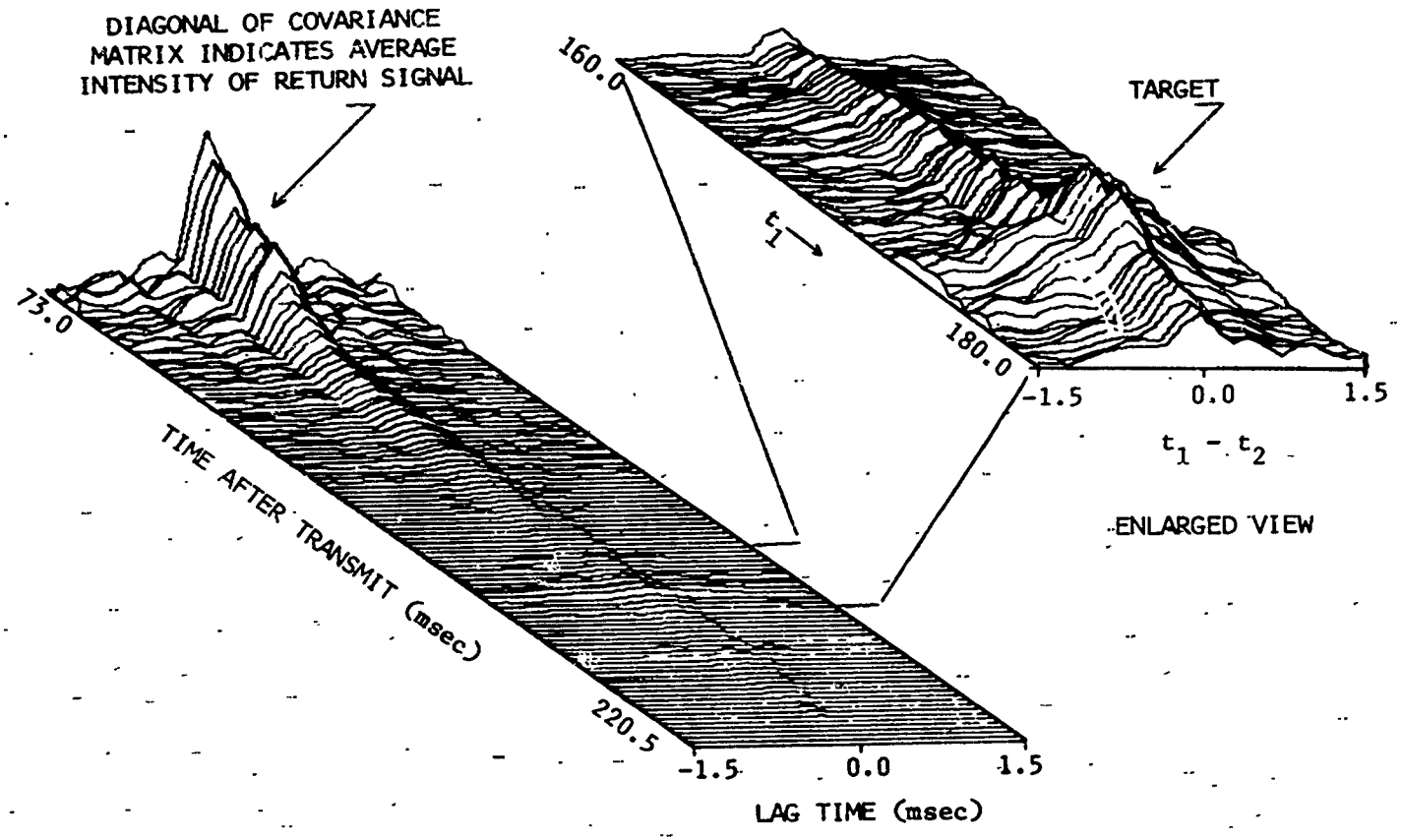


FIGURE 17

ENVELOPE OF THE COVARIANCE MATRIX ESTIMATED FROM 200 PINGS
LTTS FM DATA REVERBERATION PLUS TARGET

AS-71-535

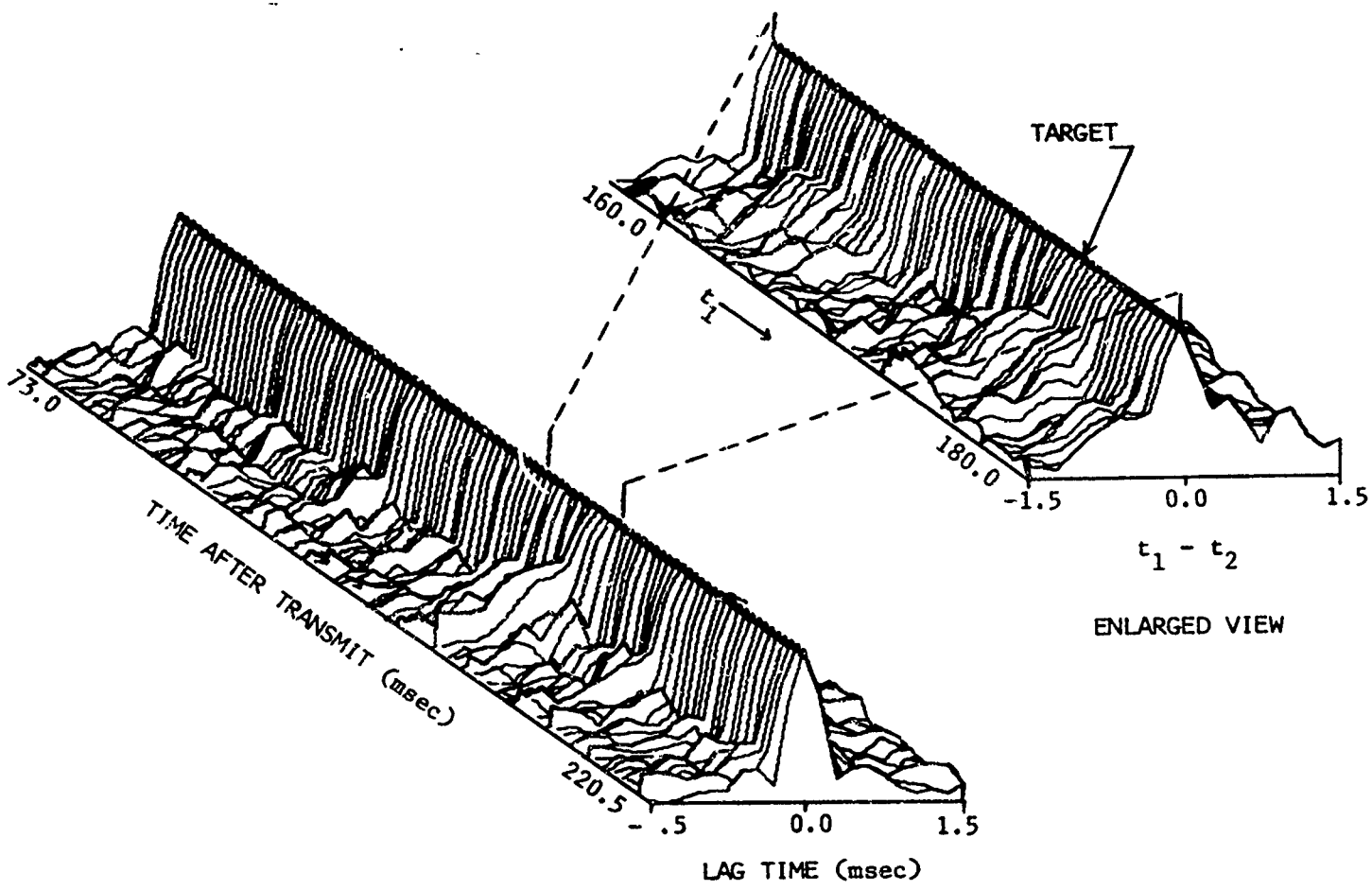


FIGURE 18

NORMALIZED ENVELOPE OF THE COVARIANCE MATRIX ESTIMATED FROM 200 PINGS
LTTS FM DATA REVERBERATION PLUS TARGET

lesser effect on the target peak. The number of cepstrums that may be averaged is limited by the relative stabilities of the sonar apparatus and the target, since a fluctuation in the relative phase (changing target aspect) could tend to cause the target peak to average to zero.

Estimated Energy Spectrum

The estimate of the backscattered sound energy spectrum $\langle X^{*}X \rangle$ is computed from the epoch times, 69 msec to 89 msec, averaged over 200 pings. The estimated energy spectrum is shown in Fig. 19, where it is observed to be shifted to the left. This is a good illustration of the significance of using the backscattered sound from a random surface, in this case the surface of Lake Travis, to flatten the energy spectrum by a division process. That is, $\langle X^{*}X \rangle$ is not uniformly constant over frequency and this would cause a significant peak to appear in the cepstrum, which would be in error. Superimposed on the estimated energy spectrum is the combined frequency response of the projector/receiver transducer system. The frequency response of the recording/reproduction system is not shown.

Detection

Once the frequency response of the entire system is known and used to flatten the energy spectrum, then any consistent peaks in the cepstrum will indicate the presence of a target. For example, Fig. 20 illustrates a typical ping return divided into 15 overlapping receive gates. A cepstrum is computed for each gate. However, the presence of a target is not apparent. The cepstrums are also shown in Fig. 21. The experiment is repeated by using three ping returns and averaging the x,y components of the cepstrums

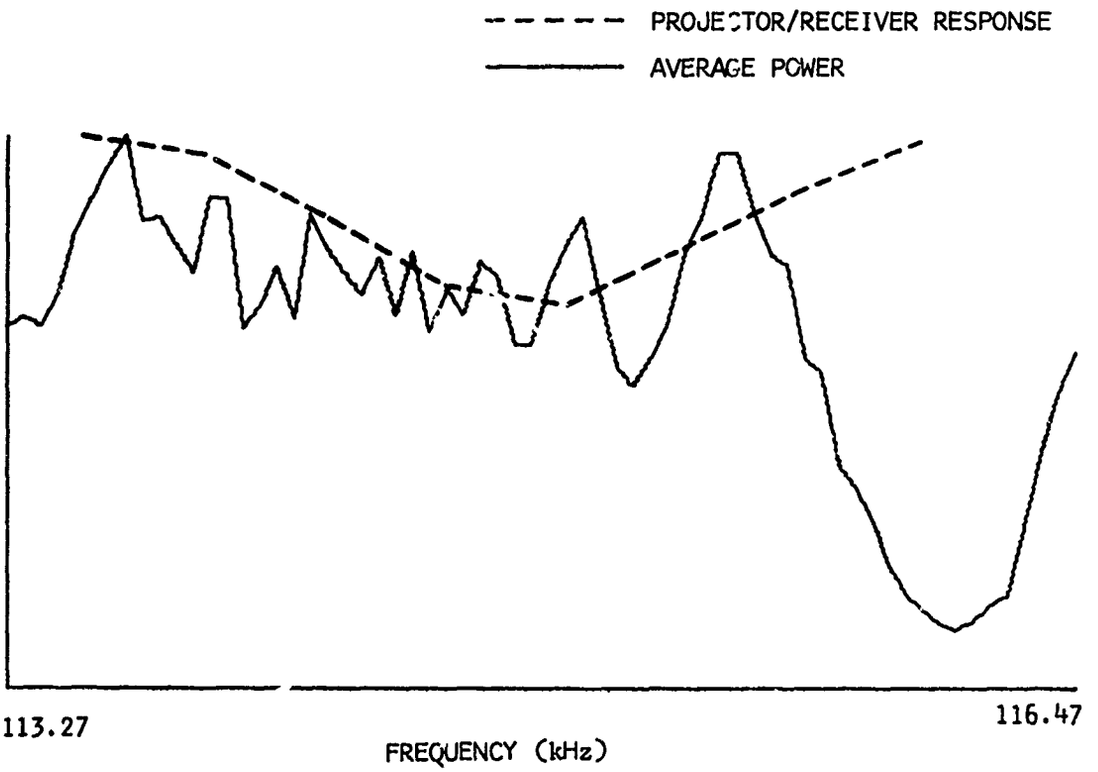


FIGURE 19

POWER SPECTRUM FROM GATE 1 AVERAGED
OVER 200 PINGS AND PROJECTOR/RECEIVER FREQUENCY RESPONSE

AS-71-537

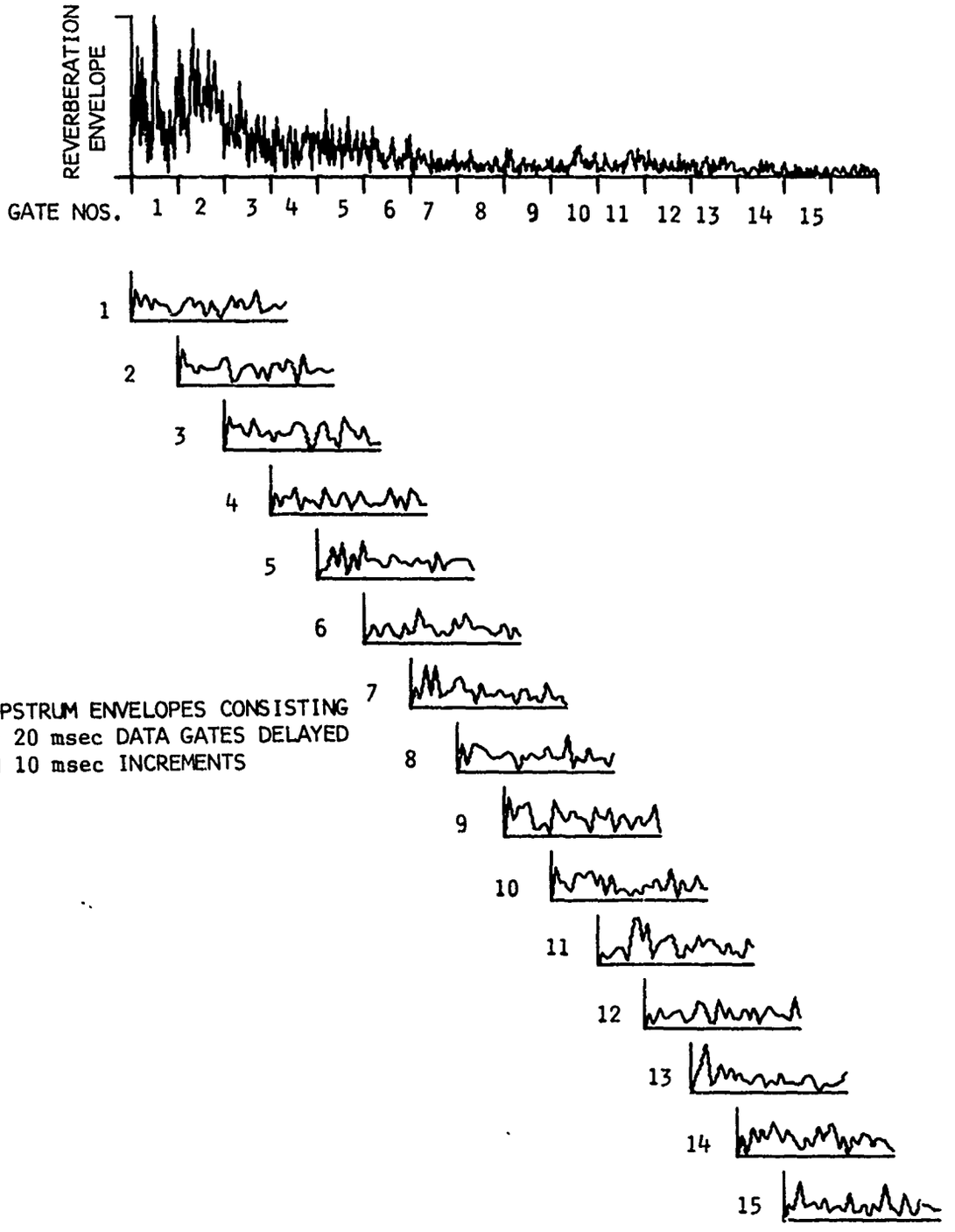


FIGURE 20

THE USE OF CEPSTRUM PROCESSING AS A DETECTOR

AS-71-538

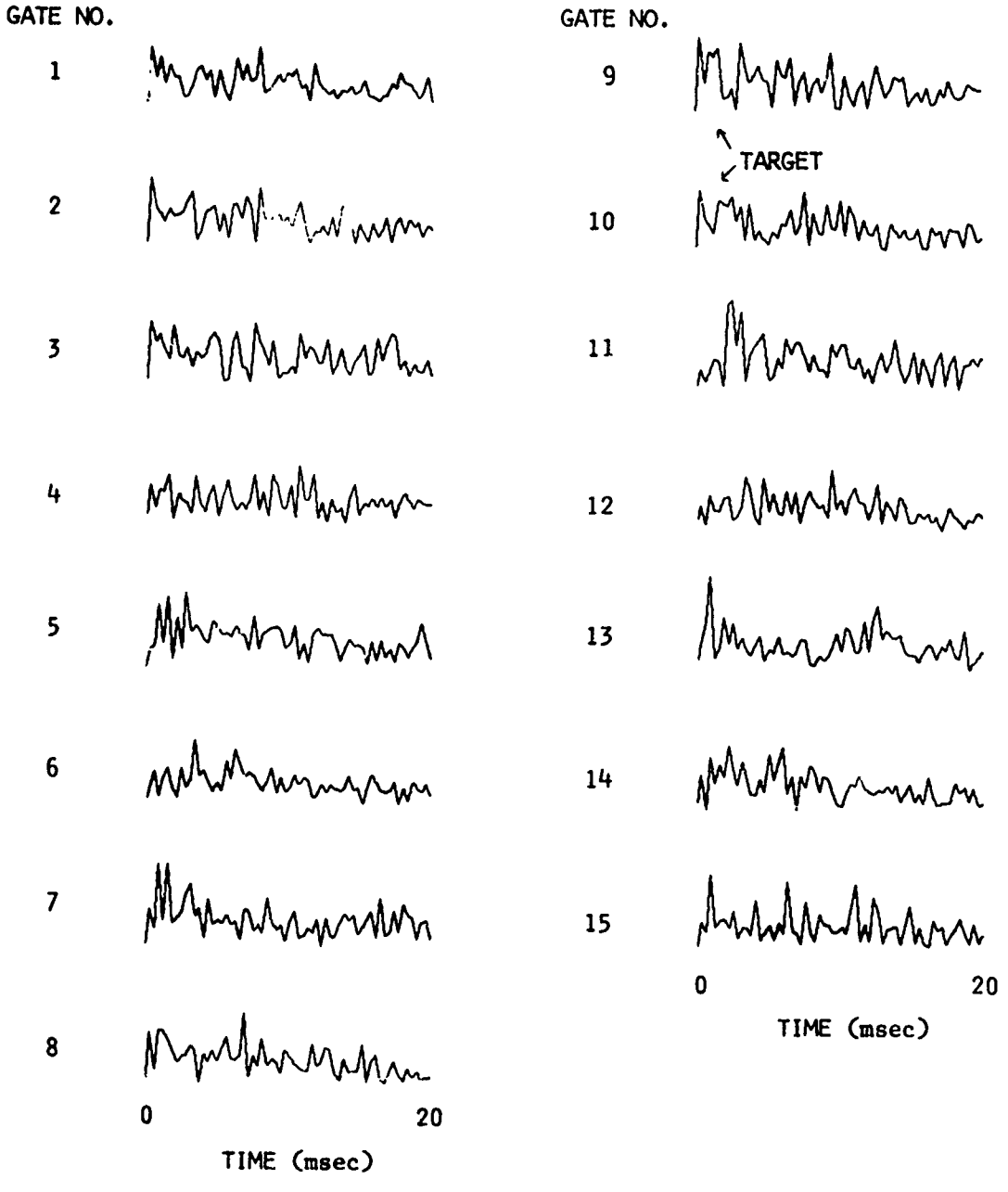


FIGURE 21

CEPSTRUMS OF LAGGED DATA WINDOWS
AVERAGED OVER ONE PING

AS-71-539

in each gate before computing the cepstrum envelopes. This time, significant peaks appear near the cepstrum origin in gates 9 and 10, as shown in Fig. 22. The experiment is repeated again by averaging over ten pings and the results are shown in Fig. 23. Again significant peaks appear close to the cepstrum origin. Also, the width of the peaks indicate an acoustic target length of 1 msec. The greatest possible acoustic length of the model target was actually 2.8 msec in this experiment. The probability of detection is illustrated by the receiver operating characteristics (ROC) shown in Fig. 24, which also shows the effects of averaging. For the purpose of generating Fig. 24, all 200 pings were used. The second, third, and fourth points of each cepstrum were used to generate histograms of amplitude, and knowing the target to be in gate 10 (gates 9 and 11 were ignored), it was possible to generate the ROC curves. The histograms from gates 1 to 8 and 12 to 15 were averaged together. The dip in the ROC curve showing an average of 10 cepstrums is probably due to an insufficient amount of data. The significant result is that averaging cepstrums greatly improves the ROC curves as illustrated by the increasing steepness of the curves on the left-hand portion of the graph.

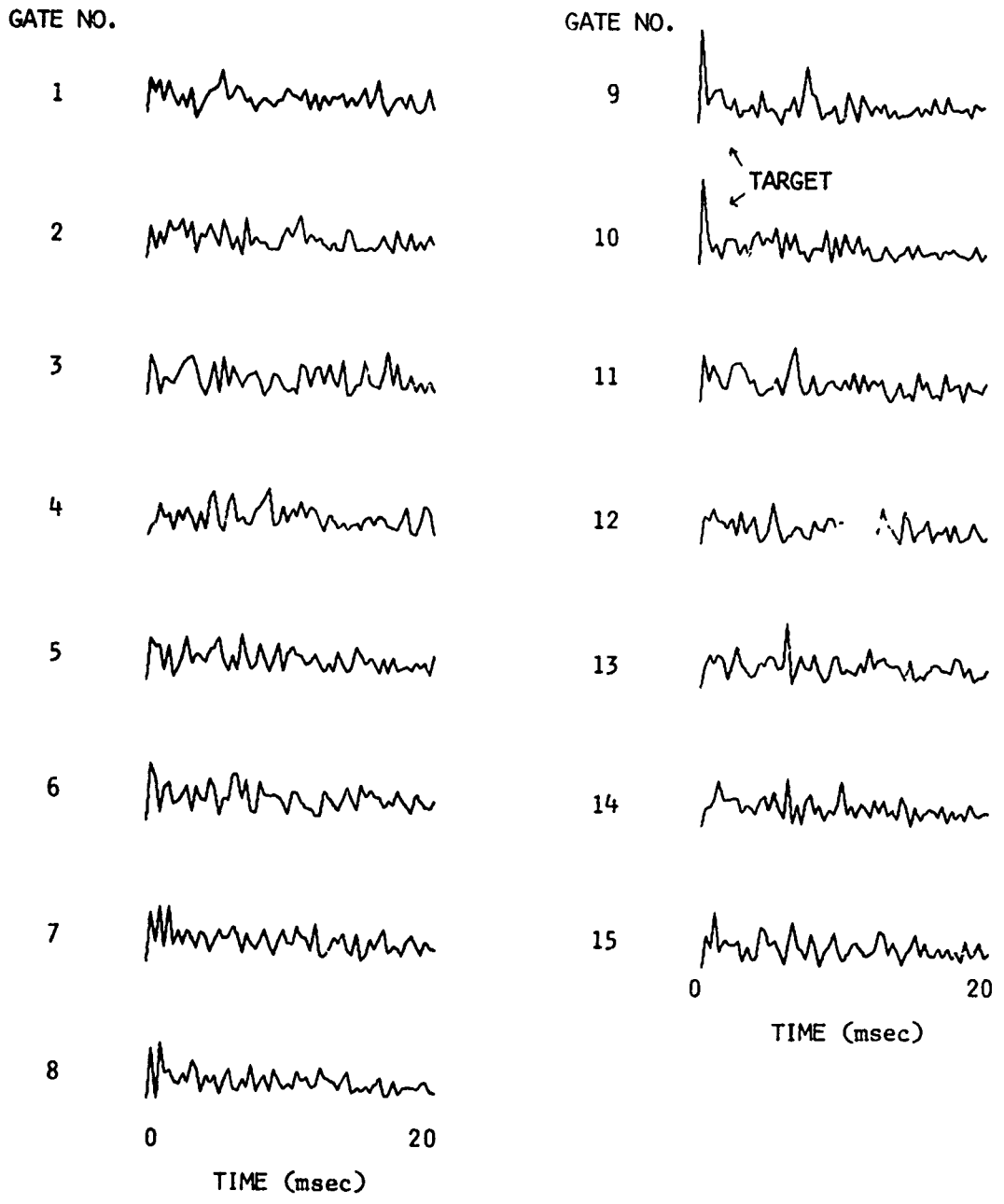


FIGURE 22

CEPSTRUMS OF LAGGED DATA WINDOWS
AVERAGED OVER THREE PINGS

AS-71-540

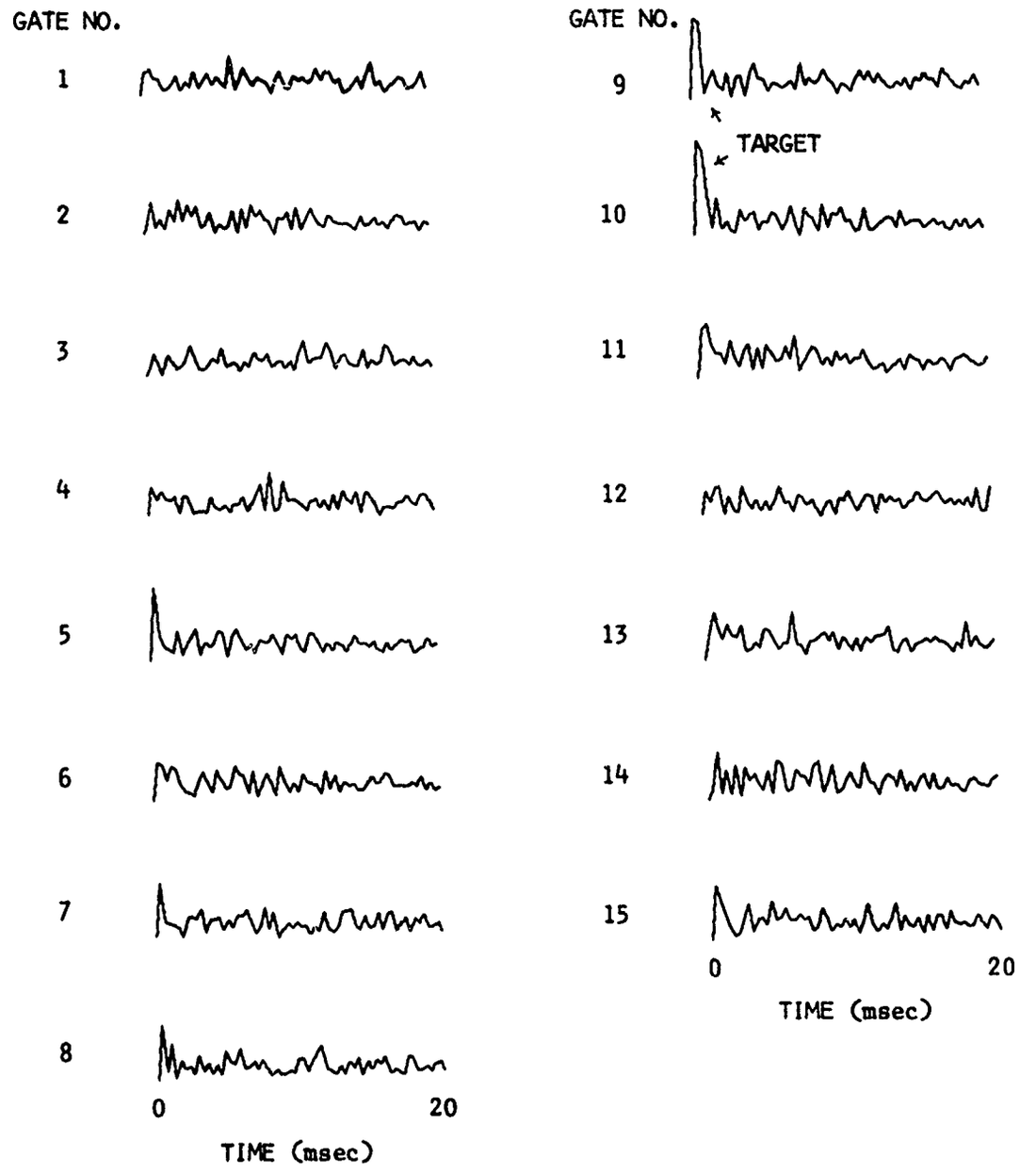


FIGURE 23

CEPSTRUMS OF LAGGED DATA WINDOWS
AVERAGED OVER TEN PINGS

AS-71-541

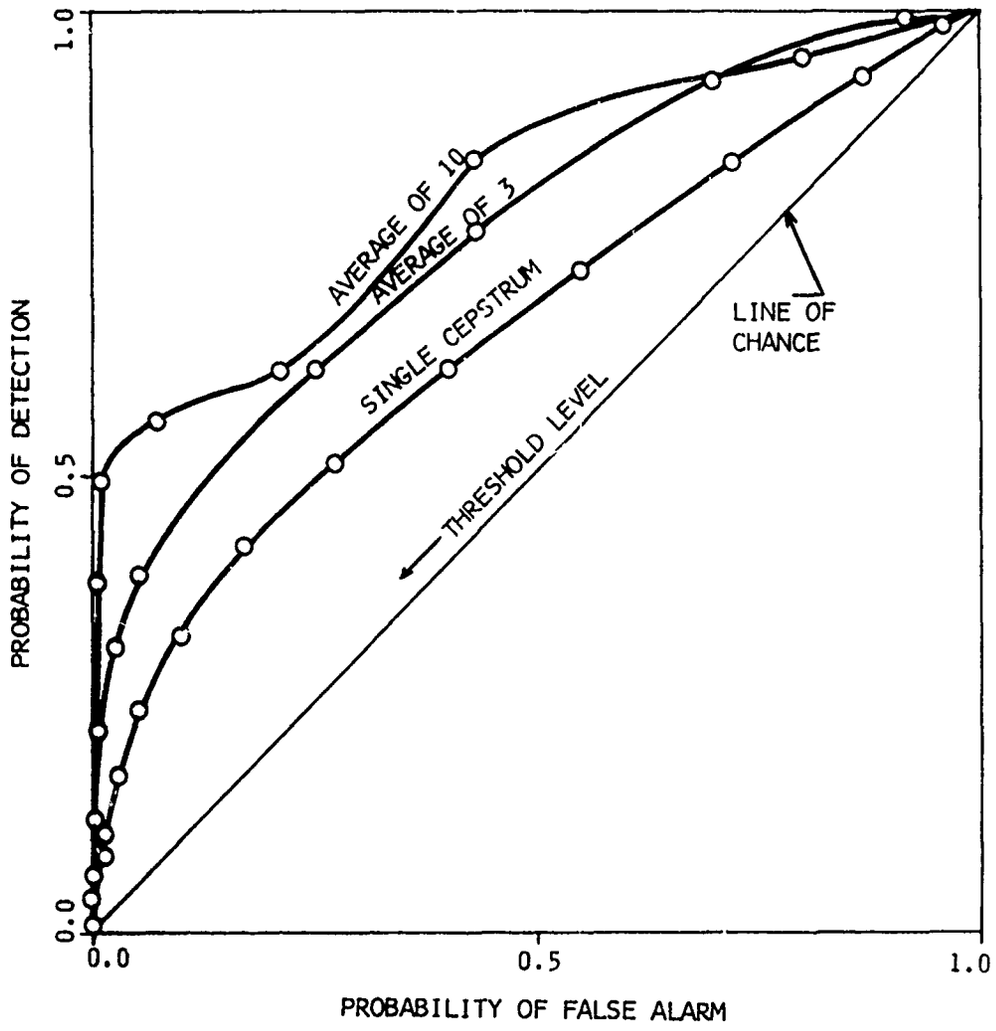


FIGURE 24

RECEIVER-OPERATOR CHARACTERISTICS OF CEPSTRUM AS A DETECTOR USING LTTS REVERBERATION PLUS TARGET DATA $(S+N)/N \approx 1.5$ dB

ARL - UT
AS-71-513
RES - DR
4 - 1 - 71

VI. FEASIBILITY OF REAL TIME CEPSTRUM

A processing system consisting of, for example, a Control Data Corporation ALPHA-1[21] computer system and a compatible Control Data FFT peripheral processing unit is considered for use as a real time cepstrum processor for a typical sonar system. The transmit signal is a frequency modulated pulse with a bandwidth of 625 Hz, carrier frequency of 5 kHz, and a pulselength of 51.2 msec (TW=32). The receive gate ranges from 606.67 msec to 12906.67 msec after the transmit; thus the maximum repetition rate is approximately one every 13 sec. The receive gate mentioned covers the range from 500 yd to 10742 yd. If the data is sampled at the bandwidth of the transmit signal, 7680 samples are necessary to estimate the signal. The length of a target is assumed to be 300 ft; thus the target length in the signal is approximately 175 msec, which is covered amply by 256 samples or a 409.6 msec data window. Fifty-nine cepstrums are necessary to cover the signal if the data window is lagged 204.8 msec or 128 samples for each cepstrum. The time necessary to process a single cepstrum is approximately 64 msec plus any small amount of time which was unforeseen in the data storage. Therefore, it should be possible to handle six signals (beams) simultaneously. The breakdown of this time is as follows:

PROCESS	TIME (msec)
1. FFT	3.086
2. Form power	10.438
3. Calculate logarithm	38.400
4. Calculate bias	1.553

Preceding page blank

PROCESS	TIME (msec)
5. Extract bias and zero Y array	2.560
6. FFT	3.086
7. Form 128 points of envelope	5.248
Total time for one cepstrum	64.371

The logarithmic step is especially time consuming, and it is assumed that additional special purpose hardware would make the processing even faster.

VII. SUMMARY

Advantages of Cepstrum Processing

The primary advantage of cepstrum is that if the data window covers a target, then the cepstrum is independent of epoch time. That is, the peaks of the cepstrum will always be located in the same relative position no matter where the target is located in the data window. The significance of this is that the quadrature components of several independent cepstrums may be averaged to achieve a processing gain. This same advantage will also be true for autocovariance and pseudoautocovariance. This advantage is not true, for example, in the case of replica correlation or matched filtering where it may be possible to average envelopes but insufficient phase control (epoch) would cause the target to average to zero if the quadrature components were used. The second advantage of cepstrum is that its mean square output is independent of the mean square input. Therefore, it is easier to choose a threshold for detection or estimation of target length.

Disadvantages

The primary disadvantages of cepstrum are its complexity and low processing gain. Even though the low processing gain may be increased by averaging, a real target's relative motion will tend to smudge the cepstrum peaks over a short period of time. Therefore it would be necessary that many averages be spread over a period of several minutes. Another problem which has been implied, but so far not mentioned, is that the signal or backscattered sound from

the target must be coherent. The reason is that cepstrum is a correlation process and it is necessary that the transmitting signal which insonifies the target be intact (coherent) and not chopped up as in a RDT-type transmission.

Conclusions and Discussion

It is concluded that cepstrum is capable of estimating target length and it is capable of detecting a target in a low signal to reverberation ratio environment. However, there is a restriction on the type of reverberation. The reverberation considered in this report was generated by sound backscattering from the air-water interface (surface), which was in random motion. Therefore, the surface reverberation will cancel in the averaging process while the backscattered sound from the target (assuming little change in aspect) will remain relatively constant. In the case of reverberation from the sea bottom it is felt that the bottom reverberation will have the same characteristics as the target in that it gives consistent returns that will not average to zero. In general, it is felt that the complexity of computing a cepstrum and the necessity of averaging over a period of several minutes outweigh the advantages of cepstrum being independent of epoch time.

REFERENCES

1. J. F. Hoffman, "Target Classification Using Likelihood Ratios and Quadrature Components," Doctoral Thesis (Physics) presented to The University of Texas at Austin, June 1969.

2. For Matched Filtering, see for example, D. O. North, Proc. IEEE 51, 1015-1027 (1963).

3. B. P. Bogert, M. J. R. Healy, and J. W. Tukey, "The Quefrency Alanysis of Time Series for Echoes: Cepstrum, Pseudo-Autocovariance, Cross Cepstrum and Saphe Cracking," in Time Series Analysis, M. Rosenblatt, Ed. (New York: John Wiley & Sons, Inc., 1963) Chapter 15.

4. A. M. Noll, J. Acoust. Soc. Am. 36, 296-302 (1964).

5. J. L. Meek, "A Note on the Mathematics of Cepstrum and an Alternate," Defense Research Laboratory Technical Memorandum No. 68-32 (DRL-TM-68-32), December 1968.

6. A. V. Oppenheim, R. W. Schafer, and T. G. Stockham, Jr., Proc. IEEE 56, 1264-1291 (1968).

7. For an example of Trend Removal, see Quarterly Progress Report No. 4 under Contract N00024-70-C-1247 (U), Systems Analysis, Applied Research Laboratories, The University of Texas at Austin. Also, see Reference No. 7.

8. V. V. Ol'shevskii, Characteristics of Sea Reverberation (New York: Consultants Bureau, 1967).

9. T. D. Plemons, "Spectra, Covariance Functions, and Associated Statistics of Underwater Acoustic Scattering from Lake Surfaces," Doctoral Thesis presented to The University of Texas at Austin, June 1971.

10. O. D. Grace and S. P. Pitt, "Quadrature Sampling of High Frequency Waveforms," J. Acoust. Soc. Am. 44, 1453-1454 (November 1968).

11. O. D. Grace, S. P. Pitt, and J. A. Shooter, "Two Applications of the Fast Fourier Transform to Aperiodic Bandpass Signals," Applied Research Laboratories Technical Report No. 69-27 (ARL-TR-69-27), 28 August 1969.

12. IEEE Transactions on Audio and Electroacoustics, Special Issue on Fast Fourier Transform and Its Application to Digital Filtering and Spectral Analysis AU-15(2), June 1967.

13. C. J. Webb, "Practical Use of the Fast Fourier Transform (FFT) Algorithm in Time Series Analysis," Applied Research Laboratories Technical Report No. 70-22 (ARL-TR-70-22), June 1970.

14. See the section on Cospectra, Reference 10.

15. J. L. Stewart and E. C. Westerfield, "A Theory of Active Sonar Detection," Proc. IRE 44, 872-881 (May 1959).

16. H. Cramer, Mathematical Statistics (Princeton: Princeton University Press, 1946), Section 18.1.

17. D. Middleton, "Acoustic Modeling, Simulation, and Analysis of Complex Underwater Targets, II. Statistical Evaluation of Experimental Data," Applied Research Laboratories Technical Report No. 69-22 (ARL-TR-69-22), The University of Texas at Austin, 26 June 1969.

18. T. D. Plemons, J. A. Shooter, and D. Middleton, "Spectra, Covariance Functions, and Associated Statistics of Underwater Acoustic Scattering from Lake Surfaces," Paper presented at the 80th Meeting of the Acoustical Society of America in Houston, Texas, 4 November 1970.

19. S. Siegel, Nonparametric Statistics, (New York: McGraw-Hill Book Co., Inc., 1956).

20. D. Middleton, IEEE Trans. Information Theory IT-13, 372-414 (1967); especially see section 9, page 402.

21. C. R. Frost, "Military CPU's" Datamation®, 87-103 (July 1970).